IMPERIAL AGRICULTURAL
RESEARCH INSTITUTE, NEW DELHI.

19336

MGIPC—S4—III-1-93—22-8-45—5,000.

PROCEEDINGS

OF THE

NATIONAL ACADEMY OF SCIENCES

OF THE

UNITED STATES OF AMERICA

VOLUME 30, 1944

EXECUTIVE COMMITTEE

OF THE PROCEEDINGS

EDWIN BIDWELL WILSON, *Managing Editor*

R. A. MILLIKAN R. G. HARRISON

PUBLISHED MONTHLY

Publication Office: Mack Printing Company, Easton, Pa.

Editorial Office: Harvard School of Public Health, Boston 17

Home Office of the Academy: Washington, D. C.

CONTENTS

MATHEMATICS

Page
Errata .... *By G. W. Mackey and H. S. Uhler* 24
Relative Number of Non-Invariant Operators in a Group .. *By G. A. Miller* 25
Note on the Function $F(a, b; c - n; z)$ .... *By H. Bateman* 28
Errata .... *By H. Bateman* 44
Birkhoff's Theory of Gravitation and Einstein's Theory for Weak Fields. *By A. Barajas* 54
On the Theory of Characters of Groups and Semi-Groups in Normed Vector Rings .... *By E. Hille* 58
The Matrices of Factor Analysis .... *By A. A. Albert* 90
Laplace Transforms of a Class of Higher Dimensional Varieties in a Projective Space of n Dimensions .... *By S. Chern* 95
Permutable Periodic Transformations .... *By P. A. Smith* 105
On the Arithmetic in a Group Ring .... *By R. Brauer* 109
Possible Numbers of Non-Invariant Operators of a Group .. *By G. A. Miller* 114
Remark on a Distortion Tensor for Elastic Displacements .... *By T. Y. Thomas* 140
The Minimum Rank of a Correlation Matrix .... *By A. A. Albert* 144
Scale Curves in Conformal Maps .... *By E. Kasner and J. DeCicco* 162
The Span of the Translations of a Function in a Lebesgue Space .... *By I. E. Segal* 165
On the Solution of a General Transform .... *By I. S. Reed* 169
Scale Curves in General Cartography .... *By E. Kasner and J. DeCicco* 211
Integral Formulas for the Characteristic Classes of Sphere Bundles .... *By S. S. Chern* 269
On Groups of Automorphism of Lie Groups .... *By C. Chevalley* 274
Groups Containing Less Than Twenty-Eight Non-Invariant Operators .. *By G. A. Miller* 275
A Group-Theoretic Characterization of the General Projective Collineation Group .... *By N. S. Mendelsohn* 279
First Proof that the Mersenne Number M_{157} Is Composite ... *By H. S. Uhler* 314
Groups Involving a Small Number of Sets of Conjugate Operators .... *By G. A. Miller* 359
A Method in Rational Diophantine Analysis .... *By E. T. Bell* 355
Some Theorems in Finite Field Theory with Applications to Fermat's Last Theorem .... *By H. S. Vandiver* 362
On Trinomial Congruences and Fermat's Last Theorem .. *By H. S. Vandiver* 368

ASTRONOMY

Galactic and Extragalactic Studies, XVII. Revision of the Distances of Thirty High-Latitude Globular Clusters .... *By H. Shapley* 61
Note on the Discovery of Red Stars .... *By G. Haro* 247

PHYSICS

The Biological Effects of X-Rays as a Function of Intensity. *By A. A. Bless* 118
The Spore Discharge Mechanism of Common Ferns .... *By A. L. King* 155
Comparison of a Degenerate Form of Einstein's with Birkhoff's Theory of Gravitation .... *By H. Weyl* 205

The Compressibility of Media under Extreme Pressures.................
By F. D. Murnaghan 244
On the Stability of Two-Dimensional Parallel Flows........*By C. C. Lin* 316
Flat Space-Time and Gravitation.........................*By G. D. Birkhoff* 324
On the Integral Equations of Continuous Dynamical Systems...........
By W. H. Ingram 370
On the Theory of the Tension of an Elastic Cylinder..*By F. D. Murnaghan* 382

ENGINEERING

On Mechanical Self-Excited Oscillations.................*By N. Minorsky* 308

BOTANY

Respiratory Rates of the Shoot Tips and Maturing Tissues in Lupinus Albus and Tropaeolum Majus................*By E. Ball and E. J. Boell* 45
Antibiotic Activity of Lichens...
By P. B. Burkholder, A. W. Evans, I. McVeigh and H. K. Thornton 250
Apomixis in Guayule...*By K. Esau* 352
Visible Expression of Cytoplasmic Pattern in the Differentiation of Xylem Strands..........................*By E. W. Sinnott and R. Bloch* 387

ZOÖLOGY AND ANATOMY

The Development of Normal and Homozygous Brachy (T/T) Mouse Embryos in the Extraembryonic Coelom of the Chick..................
By S. Gluecksohn-Schoenheimer 134
Vision and Guanine Production in Fishes................*By F. B. Sumner* 285
The Identification of Guanine in Extracts of Girella Nigricans. The Specificity of Guanase...............*By G. H. Hitchings and E. A. Falco* 294

GENETICS

A Pericentric Inversion in Gasteria Resulting in Apparent Iso-Chromosomes at Meiosis..................................*By N. H. Giles, Jr.* 1
Studies on the Development of the Eye: Evidence that the Lobe^2, Lobe^4, Lobe^5 and Eyeless^2 Mutants of Drosophila Melanogaster Develop in a Manner Similar to Bar..........................*By A. G. Steinberg* 5
Mechanism of Enzymatic Adaptation in Genetically Controlled Yeast Populations..........*By S. Spiegelman, C. C. Lindegren and L. Hedgecock* 13
Invalidation of the Cytological Evidence for Reciprocal Chiasmata in the Sex Chromosome Bivalent of Male Drosophila...*By K. W. Cooper* 50
The Behavior of Oppositional Alleles in Polyploids of Trifolium Repens.
By S. S. Atwood 69
Linkage Studies of the Rat (Rattus Norvegicus) VI.....................
By W. E. Castle and H. D. King 79
Equilibrium in Genic Materials...........................*By D. F. Jones* 82
The Effect of X-Rays upon Dominant Mutation in Maize..*By L. J. Stadler* 123
X-Ray Sensitivity Changes in Meiotic Chromosomes......*By A. H. Sparrow* 147
A Specific Abnormality Associated with a Variety of Genotypes.........
By L. C. Dunn and S. Gluecksohn-Schoenheimer 173
Can Specific Mutations Be Induced by Serological Methods?...........
By A. H. Sturtevant 176
The Induction of Mutations by Antibodies.................*By S. Emerson* 179
A Reconsideration of the Mechanism of Position Effect................
By B. Ephrussi and E. Sutton 183

Materials for the Study of the Position Effect of Normal and Mutant Genes .... *By C. Stern and G. Heidenthal* 197

Genetic Linkage of Blood Types in the Rabbit .... *By P. B. Sawin, M. A. Griffith and C. A. Stuart* 217

Concerning Linkage of Waltzing in Rats .... *By M. Whittinghill* 221

Linkage of Waltzing in the Rat .... *By W. E. Castle* 226

X-Ray Inhibition of Mitosis in Relation to Chromosome Number .... *By A. Marshak and M. Bradley* 231

Experiments on Sexual Isolation in Drosophila I. Geographic Strains of Drosophila Willistoni .... *By T. Dobzhansky and E. Mayr* 238

On Spontaneous Mutation .... *By R. Goldschmidt* 297

One-Band Inversion .... *By R. Goldschmidt and A. Hannah* 299

Chromosome Associations in Tetraploid Hybrids between Prime Type 1 and Prime Type 2 in Datura Stramonium .... *By A. D. Bergner* 302

Experiments on Sexual Isolation in Drosophila. III. Geographic Strains of Drosophila Sturtevanti .... *By T. Dobzhansky* 335

Mendelian Inheritance of Adaptive Enzymes .... *By C. C. Lindegren, S. Spiegelman and G. Lindegren* 346

Experiments on Sexual Isolation in Drosophila. II. Geographic Strains of Drosophila Prosaltans .... *By T. Dobzhansky and G. Streisinger* 340

The Synapsis of the Sex Chromosomes of Drosophila Miranda in Relation to Their Directed Segregation ... *By R. H. MacKnight and K. W. Cooper* 384

PHYSIOLOGY AND BIOCHEMISTRY

Indole and Serine in the Biosynthesis and Breakdown of Tryptophane ... *By E. L. Tatum and D. Bonner* 30

Activity and Nutritional Deprivation .... *By G. Wald and B. Jackson* 255

The Mechanics of Bioluminescence ...... *By W. D. McElroy and R. Ballentine* 377

PATHOLOGY AND BACTERIOLOGY

A Second Approximation to Soper's Epidemic Curve .... *By E. B. Wilson and J. Worcester* 37

On the Rôle of Carbon Dioxide in the Metabolism of Clostridium Thermoaceticum .... *By H. A. Barker* 88

Strain Specificity and Production of Antibiotic Substances, III. Penicillium Notatum—Chrysogenum Group .... *By S. A. Waksman and H. C. Reilly* 99

The Epidemic Curve with No Accession of Susceptibles .... *By E. B. Wilson and J. Worcester* 264

A Growth-Delaying Effect of Ultraviolet Radiation ...... *By S. E. Luria* 392

Incidence of Metabolic Changes among Virus-Resistant Mutants of a Bacterial Strain .... *By E. H. Anderson* 396

X-Ray Induced Growth Factor Requirements in Bacteria .... *By C. H. Gray and E. L. Tatum* 403

PSYCHOLOGY AND ANTHROPOLOGY

Graphical Method of Factoring the Correlation Matrix .... *By L. L. Thurstone* 129

PROCEEDINGS
OF THE
NATIONAL ACADEMY OF SCIENCES

Volume 30 January 15, 1944 Number 1

Copyright 1944 by the National Academy of Sciences

A PERICENTRIC INVERSION IN GASTERIA RESULTING IN APPARENT ISO-CHROMOSOMES AT MEIOSIS

BY NORMAN H. GILES, JR.

OSBORN BOTANICAL LABORATORY, YALE UNIVERSITY

Communicated December 20, 1943

In a study of chromosome behavior in a plant of *Gasteria* it was found[1] that an anomalous type of disjunction occurs at meiosis in one of the seven tetrad chromosomes. The chromosome involved (designated the *SML* chromosome) has a sub-median centromere. The dyads resulting from the *SML* tetrad disjoin at anaphase I in three ways:

1. *Type I.*—The two dyads, one at each pole of the spindle figure, are composed of two long and two short arms. These arms may be joined with one another as normal chromatids, or similar arms may be joined to form either long or short metacentric chromatids.
2. *Type II.*—The dyad at one pole has one short and three long arms, the dyad at the opposite pole has one long and three short arms.
3. *Type III.*—One dyad has four short arms, the other, four long arms.

The separation of the dyad chromatids at anaphase II gives rise to five different kinds of quartets with respect to the types of chromosomes derived from the original *SML* tetrad and their arrangement. These quartet types are as follows:

1. Type A (derived from Type I) has four normal *SML* chromosomes arranged in the regular manner at the poles.
2. Type B (from Type II) contains two normal chromosomes and both kinds of metacentric chromosomes, with a metacentric chromosome separating from a normal in each anaphase figure.
3. Type C (from Type III) contains only metacentric chromosomes, with the two long chromosomes separating from one another in one anaphase figure, the two short chromosomes separating from one another in the other anaphase figure.
4. Type D (from Type I) also has only metacentric chromosomes, but the two dissimilar kinds are separating from one another in each anaphase figure.

5. Type E (from Type I) is similar to Type B in having three kinds of chromosomes, but the two normal chromosomes are separating from one another in one anaphase figure, while the two dissimilar metacentric chromosomes are separating from one another in the sister anaphase figure.

The other six tetrad chromosomes of the complement behave in a regular manner at both meiotic divisions. Consequently, three kinds of microspores are produced at meiosis, all having the same chromosome number, but differing in the kind of chromosome received from the division of the aberrant tetrad. About half the microspores receive a normal chromosome from this tetrad; the other half receive metacentric chromosomes, the two complementary types being equally frequent. Only two of these microspore types, the one containing a normal haploid complement and the one having the long-armed metacentric chromosome, are able to develop and pass through the first post-meiotic mitosis, and it is clear that the development of the latter type is delayed as compared with the normal. Microspores containing a short-armed metacentric chromosome have never been observed at the microspore mitosis.

The peculiar chromosome behavior at meiosis was originally explained as resulting from a postulated aberrant type of centromere division at metaphase I of meiosis. However, it now appears unnecessary to invoke such aberrant centromere behavior in order to explain the observations. A more detailed study of meiotic prophase stages and a further analysis of the types and frequencies of the various quartets at anaphase II have shown that the results can be explained on the basis of crossing-over in a heterozygous pericentric inversion in the *SML* tetrad, a possibility suggested to the writer by Dr. Barbara McClintock.

It has not been possible to make a complete analysis of prophase pairing in the *SML* chromosome. However, synaptic figures have been found at pachytene indicating the presence of an inversion of some length, although it is not possible to be certain that the *SML* chromosome is involved. It seems probable that the figures represent a pericentric inversion since bridge configurations are very infrequent at anaphases I and II, suggesting that paracentric inversions are at most few in number and quite short. An examination of tetrad configurations has been made at diplotene and diakinesis. It is very difficult to fix these so-called diffuse stages properly in *Gasteria*, but reasonably good results have been obtained by using an osmic acid fixative, staining by the Feulgen method, and counterstaining with fast green to reveal the nucleolus.[2] Since the nuclelous is attached to the long arm of the *SML* tetrad, the latter can be easily located. Some of the configurations of this tetrad chromosome strongly suggest by the position of chiasmata and the relationships of the long and short arms that crossing-over has occurred within an inversion including the centromere.

The analysis of quartets at AII is particularly instructive, since the various types and their frequencies can all be logically explained as a result of single or double crossing-over within a pericentric inversion whose length is approximately equal on each side of the centromere. The synapsis of a normal chromosome and one with the inversion would give rise to an inversion loop configuration with the centromere in approximately the middle of the loop. Crossing-over within this inversion produces the previously described types of quartets as follows:

1. Type A results from (*a*) no crossing-over within the inversion, (*b*) two-strand double crossing-over on one side of the centromeres, either to the right or to the left, or (*c*) two-strand double crossing-over, the two crossovers being on opposite sides of the centromere.
2. Type B results from (*a*) a single crossover on either side of the centromeres, or (*b*) a three-strand double on one side, either to the right or to the left.
3. Type C results from a four-strand double on one side of the centromeres, either to the right or to the left.
4. Type D results from a four-strand double, the two crossovers being on opposite sides of the centromeres.
5. Type E results from a three-strand double, the two crossovers also being on opposite sides of the centromeres.

TABLE 1

FREQUENCIES OF THE FIVE TYPES OF QUARTETS AT AII. FOR A DESCRIPTION OF THE TYPES AND A DISCUSSION OF HOW THEY ARISE SEE TEXT

TYPE OF QUARTET	NO. OF CELLS	PER CENT
A	126	18.7
B	374	55.4
C	57	8.4
D	40	5.9
E	78	11.6
Total	675	100.0

The frequencies of the five types are indicated in table 1. As is expected, the types (C, D and E) which result from double crossing-over only are much less frequent than Type B, which is the only single crossover type. A further comparison can be made between Types D and E which both result from double crossing-over, the two crossovers being on opposite sides of the centromere. Type D results from a four-strand double and Type E from a three-strand double. If crossing-over involves the four chromatids at random[3, 4], E should be twice as frequent as D. It is clear from table 1 that E is indeed approximately twice as frequent as D. The possible distorting effect on this ratio of triple crossing-over (with two crossovers on one side of the centromeres and one on the opposite side) has been considered. It is found that triple crossing-over within the inversion does not affect the D:E ratio, since with random chromatid crossing-over Types B, D and E are produced, but the D:E ratio remains 1:2. Thus there is good evidence for random crossing-over of chromatid strands. The comparative frequencies of the various quartet types resulting from double crossing-over (especially Types C and D) indicate that double crossovers on one side of the centromeres are higher than double crossovers involving segments on opposite sides of the centromeres. The reason for this result is not clear, but it may mean that incomplete synapsis occurs on one side of the inversion loop in some of the pollen mother cells. Synapsis of this type should lead to an increased frequency of double crossovers on one side of the centromeres as compared with doubles involving segments on opposite sides.

The presence of chromosomes with a median centromere at anaphase II indicates that the inversion in the altered *SML* chromosome must be of approximately equal length on either side of the centromere. Otherwise, chromosomes with non-median centromeres would result from crossing-over. Furthermore, it is clear that crossing-over must occur with a very high frequency in the proximal regions of the *SML* tetrad (and observations on chiasmata at diplotene support this conclusion) in order to account for the frequencies of novel quartet types at AII. This unusual combination of circumstances gives rise to the configurations at anaphases I and II which were originally interpreted as iso-chromosomes produced by centromere misdivision at metaphase I. The present evidence indicates that the two types of metacentric chromosomes are not true iso-chromosomes in the sense of Darlington,[5] since they result from crossing-over and consequently do not have genetically identical arms. Morphologically, however, the metacentric chromosomes are indistinguishable from iso-chromosomes in metaphase and anaphase figures. Furthermore, they are also similar to iso-chromosomes in being duplication-deficiency types. The fact that the metacentric chromosomes carry certain regions of the *SML* chromosomes in duplicate and are deficient for other regions explains the previously mentioned observations made at the first post-meiotic mitosis in the developing microspores. It will be recalled that microspores carrying the short-armed metacentric chromosome (having a long deficiency) never develop as far as the pollen mitosis whereas those microspores carrying the long-armed type (having a short deficiency) do go through the pollen mitosis, but are delayed in their development as compared with microspores having a normal chromosome complement.

Summary.—Further cytological studies have been made of a plant of *Gasteria* which gives rise to large numbers of two types of metacentric chromosomes (in addition to the normal type) as a result of the aberrant behavior of a single one of the seven tetrad chromosomes at meiosis. The evidence obtained from meiotic prophase configurations and from the types and frequencies of the various quartets at anaphase II indicates that these new types of chromosomes arise from crossing-over in a heterozygous pericentric inversion rather than from aberrant centromere behavior as previously suggested. The presence of metacentric chromosomes indicates that the inversion is of approximately equal length in the two arms of the altered chromosome and the frequencies of novel quartet types show that crossing-over occurs within the inversion with a high frequency. The ratio of comparable quartet types resulting from three- and four-strand double crossing-over agrees with that expected if crossing-over involves the four chromatids at random.

---

[1] Giles, Norman H., Jr., *Genetics*, **28**, 512–524 (1943).

[2] Hillary, B. B., *Bot. Gaz.*, **102**, 225–235 (1940).

[3] Beadle, G. W., and Emerson, S., *Genetics*, **20**, 192–206 (1935).

[4] Weinstein, A., *Ibid.*, **21**, 155–199 (1936).

[5] Darlington, C. D., *Jour. Genet.*, **39**, 351–361 (1940).

---

*STUDIES ON THE DEVELOPMENT OF THE EYE: EVIDENCE THAT THE LOBE^2, LOBE^4, LOBE^5 AND EYELESS^2 MUTANTS OF DROSOPHILA MELANOGASTER DEVELOP IN A MANNER SIMILAR TO BAR***

BY ARTHUR G. STEINBERG

DEPARTMENT OF GENETICS, McGILL UNIVERSITY

Communicated December 13, 1943

Introduction.—Studies in the field of developmental genetics, although primarily directed toward an understanding of how the gene acts, have in many instances led toward a better understanding of the developmental mechanics of the organisms studied. This is most certainly true of *Drosophila melanogaster*, in which our knowledge of its embryology, both descriptive and experimental, has been derived from studies concerning the development of various mutant forms.

Studies on the development of the Bar mutation of *D. melanogaster* have led to a hypothesis concerning its (the mutation's) mode of development which, if capable of generalization, would add further information to our growing knowledge of the embryology of Drosophila. These studies showed that the Bar eye disc was smaller than that of wild type at all larval stages from 36 hours after hatching, the earliest stage studied, to pupation. They demonstrated that the rate of increase in size of the Bar eye disc was the same as that of wild type and that histological differentiation proceeded at the same rate as in wild type (Steinberg 1941*a*, 1941*b* and 1943). They have also established that both a second chromosome inhibitor of Bar, and temperature bring about their effect on facet number without affecting the size of the eye disc, although it had been shown by Medvedev (1935) for Lobe,^{c} glass^{2}, eyeless^{2} and wild type, and by Steinberg and Abramowitz (1938) for double-Bar, double infra-Bar, Bar, infra-Bar and Bar^{4} that the size of the eye discs of the mature larva is directly proportional to the number of facets in the imaginal eye.

The hypothesis based on these studies postulates that the reduced size of the Bar eye disc is due to the participation of fewer cells in the initial formation of the eye disc, and that this smaller size of the eye disc determines the range over which the facet number of the adult eye may vary; the variation in the facet number, once the eye disc is formed, is postulated to be due to the presence of cells labilely determined to form either head chitin or facets. The ultimate fate of these cells is dependent upon various extrinsic and intrinsic factors—such as temperature, nutrition, modifiers, etc. (Steinberg 1941*a*, 1941*b* and 1943). Reduced to its bare essentials the hypothesis postulates two main steps in the determination of the facet number of Bar, the first in the embryonic stage when the size of the eye disc and hence the possible range of variation of facet number are determined, the second in the larval stage when the fate of the labilely determined cells is fixed.

Some evidence exists which indicates that at least the first step postulated in the determination of the facet number is present in other mutant forms as well as in Bar. The evidence is derived from the data of Medvedev and of Steinberg and Abramowitz cited above which showed that a reduction in the facet number of the imaginal eye is associated with a proportional reduction in the size of the eye discs of the mature larvae. Hence all of the eight mutants studied in this respect show indications of the presence of the postulated first step in the determination of the adult facet number. There is no evidence available concerning the presence or absence of labilely determined cells in any mutant type other than Bar. This paper reports the results of experiments designed to obtain such evidence for four mutants of *D. melanogaster*.

Materials and Methods.—The mutants used were the alleles Lobe^{2}, Lobe^{4} and Lobe^{5} (L^{2}, L^{4} and L^{5}, respectively) located at 72.0 in the second chromosome and eyeless^{2} (ey^{2}) located in the fourth chromosome. The stocks bearing the mutants were maintained in mass cultures.

These mutants were chosen for study because previous experience had shown that the facet number of flies of each of these genotypes is extremely variable even in isogenic stocks raised under uniform conditions, and that the facet numbers of the left and right eyes of flies of each of these mutant types are largely independent of each other. It was reasoned that this variability is due, at least in part, to an extreme sensitivity of the labilely determined cells to minor fluctuations in the environment in which the larvae develop.

If the variation in facet number from fly to fly within a given genotype is due simply to changes in the size of the eye discs, it follows that the coefficient of variability of the facet number should not be greater than that of the size of the eye discs of mature larvae of the same strain. On the other hand, if the variation in facet number from fly to fly is associated with the presence of labilely determined cells, the coefficient of variation of the facet number should exceed that for the size of the eye discs. Likewise, if the variation in facet number between the left and right eyes of a given fly is due solely to variation in the size of the left and right eye discs it follows that the left-right correlation values for the eye discs should be the same as that for the corresponding facet numbers, while, on the other hand, if the variation is due to the presence of labilely determined tissue, the left-right correlation of the eye discs should exceed that of the facets.

The experiment consisted, therefore, in counting the facet number of the left and right eyes of flies of each of the four mutant strains and of determining the areas of the left and right eye discs of mature larvae of each of the four. The facet counts were made by dissecting off the cornea of the eye, mounting it on a slide in a drop of water and then projecting the image. This method permits of great accuracy in the determination of the facet number, hence the statistics determined from these data are also accurate. The areas of the optic discs were determined by means of planimeter tracings of camera lucida drawings of eye discs newly dissected from mature larvae in a Ringer's solution (NaCl, 7.5 g.; KCl, 0.35 gms; $CaCl_2$, 0.2 g.). Because the discs are curved they must be flattened against the slide in order that the camera lucida drawings be accurate. The process of flattening, involving tapping with a needle, distorts the shape of the disc and doubtlessly introduces variations in the area of the eye discs, which may be either increases or decreases. In addition there are random errors introduced in the course of the planimeter tracings. These errors tend to increase the variability of the eye discs within a group, and precisely because they are random to decrease the left-right correlation calculated from the measurements. Consequently, the values of the statistics calculated from the eye disc measurements are shifted in such a way as to reduce their sensitivity. A further source of error is the relatively poor control of the ages of the larvae (discussed in a following paragraph). Since the eye discs increase in size as the larvae age, any variation in ages would increase the variability of the disc size from fly to fly and again reduce the sensitivity of the statistics.

The imagos for the facet counts were raised on the standard corn meal-agar-molasses food medium, reënforced with dried brewers' yeast. The larvae used for eye disc measurements were obtained as follows: 20–30 females were permitted to lay eggs on "wafers" of food (approximately 2% Carragar plus 12.5% molasses in water), over a 24-hour-period. At the end of this period the wafers were transferred to refrigerator dishes which were half filled with the standard food medium. Larvae which had crawled up the sides of the dish preparatory to puparium formation were considered mature and were the source of the eye discs measured in the various experiments.

With one exception, all experiments were done at 25 ± 0.2°C. The exception involved the Lobe^4 flies raised for facet counts when, due to a heat wave, the temperature fluctuated between 24°C. (when ice was added to the incubator) and 28°C. It was usually closer to the latter than the former.

TABLE 1

MEAN NUMBER OF FACETS IN THE LEFT AND RIGHT EYES, AND THE MEAN AREAS OF THE CAMERA LUCIDA DRAWINGS OF THE EYE DISCS OF MATURE LARVAE OF EACH OF THE FOUR MUTANT TYPES STUDIED. THE AREAS OF THE CAMERA LUCIDA DRAWINGS ARE IN SQUARE INCHES × 100. THE LINEAR MAGNIFICATION = 114.06 ×

		FACETS			EYE DISCS		
		LEFT EYE	RIGHT EYE		LEFT EYE	RIGHT EYE	
GENOTYPE	SEX	$\bar{X} \pm \sigma_{\bar{x}}$	$\bar{X} \pm \sigma_{\bar{x}}$	N	$\bar{X} \pm \sigma_{\bar{x}}$	$\bar{X} \pm \sigma_{\bar{x}}$	N
L^2	♂	106.0 ± 9.5	92.9 ± 9.1	35	48.9 ± 2.1	49.7 ± 2.0	25
	♀	134.4 ± 12.7	142.9 ± 10.5	31	51.5 ± 2.5	53.4 ± 2.2	24
L^4	♂	117.6 ± 10.7	104.0 ± 8.9	26	48.9 ± 1.7	49.0 ± 1.6	25
	♀	126.9 ± 9.3	147.2 ± 8.6	37	56.1 ± 1.9	60.5 ± 2.5	23
L^5	♂	135.8 ± 11.1	151.5 ± 9.8	26	50.2 ± 1.8	51.0 ± 2.6	25
	♀	153.8 ± 10.5	159.2 ± 11.3	22	52.6 ± 2.2	54.8 ± 2.0	25
ey^2	♂	279.8 ± 22.4	339.4 ± 22.6	19	79.0 ± 3.6	77.2 ± 3.3	25
	♀	289.1 ± 21.7	290.8 ± 22.2	16	76.6 ± 5.2	74.2 ± 4.2	25

Data.—The mean facet number of the imaginal eye and the mean areas of the camera lucida drawings of the eye discs will be found in table 1. They are presented for the purpose of orientation and as a summary of the raw data from which tables 2 and 3 have been derived.

Table 2, which presents the correlation coefficients of the facet numbers in the left and right imaginal eyes and of the larval eye disc sizes, shows that in every case r (correlation coefficient) is greater for the eye discs than it is for the facet numbers. Furthermore, of the eight comparisons only two, namely L^2 females and L^4 males, give a value of P which exceeds 0.05. It must be emphasized, however, that even in these two cases r is greater for the eye discs than for the imaginal eye. Furthermore, as pointed out in the previous section, the estimated value for r for the imaginal discs is certainly lower than its true value and therefore the observed differences between the values of r for the facet numbers and the eye disc sizes are in every case lower than the true differences.

TABLE 2

CORRELATION VALUES FOR FACET NUMBERS OF THE LEFT AND RIGHT EYES AND FOR THE AREAS OF THE LEFT AND RIGHT EYE DISCS, AND THE VALUE OF D/σ_D FOR THE COMPARISONS BETWEEN THE STATISTICS OF THE FACET NUMBERS AND THE EYE DISC AREAS

GENOTYPE	FACETS			EYE DISCS			
	SEX	r	N	SEX	r	N	D/σ_D*
L^2	♂	0.03	35	♂	0.58	25	2.25
	♀	0.41	31	♀	0.67	24	1.27
L^4	♂	0.36	26	♂	0.65	25	1.33
	♀	0.16	37	♀	0.68	23	2.39
L^5	♂	0.35	26	♂	0.74	25	1.93
	♀	0.09	22	♀	0.66	25	2.26
ey^2	♂	−0.25	19	♂	0.84	25	4.42
	♀	0.04	16	♀	0.86	25	3.75

* Since the question asked is, "is the coefficient of correlation for the eye discs *greater than* that for the facets?" only one tail of the normal curve is used, hence $P = 0.05$ at $D/\sigma_D = 1.64$ and $P = 0.01$ at $D/\sigma_D = 2.33$ (comparisons made by converting r to Z).

TABLE 3

COEFFICIENTS OF VARIATION FOR THE FACET COUNTS AND EYE DISC AREAS AND THE VALUES OF D/σ_D FOR THE COMPARISONS OF THE STATISTICS OF THE CORRESPONDING FACET NUMBERS AND EYE DISC AREAS

GENOTYPE	SEX		FACETS $CV \pm \sigma_{cv}$	EYE DISCS $CV \pm \sigma_{cv}$	D/σ_D*
L^2	♂♂	Left	52.6 ± 6.4	21.1 ± 3.2	4.4
		Right	57.9 ± 7.0	19.5 ± 3.0	5.1
	♀♀	Left	52.7 ± 6.8	23.7 ± 3.6	3.8
		Right	40.9 ± 5.3	20.6 ± 3.2	3.3
L^5	♂♂	Left	41.8 ± 5.9	18.1 ± 2.8	3.6
		Right	32.9 ± 4.7	25.9 ± 3.8	1.1
	♀♀	Left	31.9 ± 4.9	20.9 ± 3.1	1.9
		Right	33.4 ± 5.2	18.0 ± 2.8	2.5
L^4	♂♂	Left	46.4 ± 6.5	17.2 ± 2.6	4.1
		Right	43.5 ± 6.2	15.9 ± 2.5	4.2
	♀♀	Left	44.5 ± 5.3	16.6 ± 2.7	4.8
		Right	35.3 ± 4.3	20.2 ± 3.2	2.8
ey^2	♂♂	Left	34.9 ± 5.8	22.7 ± 3.4	1.8
		Right	27.8 ± 4.9	21.6 ± 3.2	1.05
	♀♀	Left	31.3 ± 5.6	34.2 ± 5.0	−0.37
		Right	30.6 ± 5.5	28.7 ± 4.2	0.28

* Since the question asked is, "is the coefficient of variation for the eye discs *less than* that for the facets?" only one tail of the normal curve is used, hence $P = 0.05$ at $D/\sigma_D = 1.64$ and $P = 0.01$ at $D/\sigma_D = 2.33$.

Table 3 contains the data pertinent to a comparison of the coefficients of variation (*Cv*) of the larval eye discs and of the facet numbers of the imaginal eyes. Of the twelve comparisons among the Lobe alleles (i.e., facet number of the left imaginal eyes against the area of left larval eye discs, and right eyes against right eye discs of the males and females separately, making four comparisons within each genotype) only the one concerned with the right eyes and discs of the Lobe^5 males results in a value of P greater than 0.05. This seems to be due to the relatively large coefficient of variation (*Cv*) shown by right eye discs of the Lobe^5 male larvae. The raw data of these eye disc measurements show one eye disc which measured 77 and another which measured 88; the next closest value to these is 61. If these two measurements are tested as possible samples of a population composed of the remaining 23 measurements it is found that neither of them fit, i.e., $P < 0.01$. Therefore, there is considerable justification for excluding these measurements from the sample. If these two values are excluded $\bar{X} = 48.3 \pm 1.64$ and $Cv = 12.2 \pm 1.8$. If this Cv is compared with that of the facet number D/σ_D is found to be 4.1. It seems valid therefore to conclude that the coefficient of variation of the eye discs of the Lobe^5 males is smaller than that for the facet number, and therefore that the Lobe^5 males conform to the same pattern as the other lobe flies involved in the experiment.

The coefficients of variation derived from the ey^2 data quite clearly do not show the same relationships that those for the various lobe alleles do (table 3). In only one case, the left eyes of the ey^2 males, is P less than 0.05. In the remaining cases it is considerably above 0.05. When the data for each of the four sets of eye discs are plotted in a scatter diagram, no tendency to cluster about a modal value is seen. The 25 values of each group are arranged in an almost unbroken sequence over the entire range with very little repetition of any given value. Thus in the data for the left eye discs of the males, five values occur twice, none more than twice; in the case of the right eye, one value occurs three times and all the others once each; in the data for the left eye disc of the females three values occur twice each, none more than twice; and, finally, only one value occurs twice in the data for the right eye discs of the females. Clearly the high values of the coefficients of variability shown by the eye discs of the ey^2 males and females are not due to one or two exceptional individuals as is the case with the right eye discs of the Lobe^5 male larvae but are a true expression of the variability of the sample. The increased variability of eyeless^2 eye discs relative to that of the eye discs of the Lobe alleles may be, in part or in whole, due to variation in the age of the larvae selected for dissection, for as pointed out above the age control was relatively poor. This problem will be returned to below.

Discussion.—The theory of development of the eye discs reviewed in the introduction of this paper postulated in part that variation between the left and right eyes of a given individual might result from a variation in the fate of labilely determined cells in the larval eye discs. Such variation would not result in a change in the size of these discs, hence the left-right correlation should under these circumstances be greater for the larval eye discs than for the facet numbers of the imaginal eyes. The data of table 2 show quite clearly that the requirements of the hypothesis are fulfilled in each of the four mutant types studied. Hence in each of these mutant types the facet number of the imaginal eye may be varied during the larval stages by altering the path of development of labilely determined cells. These cells, in the Lobe alleles and eyeless^2, must be much more sensitive to minor variations in the environment than they are in Bar and wild type since in the former Margolis (1936) has shown a significant left-right correlation for the eyes of genetically Bar flies and Margolis and Robertson (1937) have shown that even marked changes in temperature during larval development fail to cause any great shift in the facet number of wild type.

These experiments taken in conjunction with those on Bar and modified Bar (Steinberg 1941*a* and 1941*b*) indicate that mutations may influence the labilely determined cells with regard to their sensitivity to environmental changes (*ey*^{2}, the Lobe alleles and Bar), and with regard to the path in development which they will take (modifier of Bar). Presumably there are mutants (perhaps among those already studied) which influence the number of labilely determined cells present in an eye disc, but no technique for detecting such effects is available at the moment.

The data in table 3, concerning the variability of the facet number of the imaginal eyes as compared with that of the corresponding eye discs, when considered in conjunction with the data of table 2, afford evidence in regard to the manner in which the range of variation of the facet number is determined in each of the specific genotypes. Variation in the facet number of the imaginal eye may arise as the result of variation in the number of cells entering into the formation of the eye disc (detectable as a variation in the size of the discs) or as a variation in the number of labilely determined cells which finally form facets (detectable by means of a higher left-right correlation for the size of the larval eye discs than for the facet number of the imaginal eye) or through some combination of the two (detectable by a combination of the above methods). The data of table 3 show that the coefficients of variation for the eye discs are in all cases significantly smaller than those for facet number. Nevertheless, the coefficients of variation for the eye discs are rather high, particularly when compared with the values derived for Bar and wild type eye discs in the author's earlier studies. However, it is entirely likely that a large part of the relatively increased variability is due to the much less accurate control of age employed in these experiments. It will be recalled that in the present experiments age was determined simply by observing which larvae had crawled up the side of the dishes preparatory to puparium formation. In the earlier experiments, newly hatched larvae were collected over a 2-hour period and raised in 3-inch petri dishes. In the mature age group, measurements were made of the eye discs of only those larvae which were unpupated at the time when about one-half the total number of larvae in the dish had pupated. Therefore, it seems valid to conclude that in the Lobe alleles, as in Bar, the variation in facet number occurring from fly to fly is due to the variation in the number of labilely determined cells which go to form facets rather than head chitin.

The story seems to be quite different with eyeless^2 for here there is no difference between the coefficients of variation of the eye discs and of the facet numbers. Comparison of the data for the eye discs of eyeless^2 larvae with those for the eye discs of the Lobe alleles (table 3) shows that the values of the coefficients of variation of the latter are lower than those of the former. A similar comparison of the data for the facets shows that while the coefficients of variation for the facet numbers of Lobe^2 and Lobe^4 eyes are larger than those for eyeless^2, those for Lobe^5 are about the same. There is, then, a disproportionate increase in the variability of the areas of the eye discs of ey^2 larvae. There can be no doubt that the values of the coefficients of variation derived from the facet numbers of imaginal eyes are accurate since the facet counts are accurate. The same is certainly not true of the values derived from the eye disc data, for, as pointed out in the introduction, the eye discs vary in size with age and any inaccuracies in age determination would tend to increase the coefficients of variability and superimposed on this variation is the variation introduced by the technique of measurement. It is possible, therefore, that the absence of a difference between the coefficients of variation of the eye discs and of the facets is more apparent than real. Only further experiments can settle this question. If the data are correct, it would mean that the variation in facet number from fly to fly in eyeless^2 is dependent upon variation in the number of cells entering into the formation of the eye discs. Superimposed upon this would be the variation in the number of labilely determined cells which go to form facets. It is this variation which affects the left-right correlation coefficients. It is unlikely that variation in the number of cells entering into the formation of the eye discs would influence the value of r because from the nature of the development of the eye discs it is highly probable that equal numbers of cells would enter into the formation of each disc.

Summary.—Previous experiments had led to the hypothesis that the facet number of the Bar eye of *Drosophila melanogaster* was determined as follows: (*a*) a reduced number of cells enter into the formation of the eye discs thus limiting the range over which the facet number of Bar eyes can vary and (*b*) a group of labilely determined cells which may form either head chitin or facets depending upon the nature of the extrinsic and intrinsic environment determine where, within the possible range of variation of facet number, the facet number of a given eye will fall. The experiments reported in this paper were designed to test the applicability of this hypothesis to the development of Lobe^2, Lobe^4, Lobe^5 and eyeless^2. The data show that the development of these mutants may be adequately described by the theory.

* Aided by a grant from the Penrose Fund of the American Philosophical Society.

Margolis, O. S., "Studies on the Bar Series of Drosophila III. The Facet Relations in Dorsal and Ventral Lobes of the Eye," *Genetics*, **21**, 817–831 (1936).

Margolis, O. S., and C. W. Robertson, "Studies on the Bar Series of Drosophila IV. The Temperature Effective Period for Facet Determination in the Wild Type," *Genetics*, **22**, 319–331 (1937).

Medvedev, N. N., "Genes and Development of Characters. I. The Study of the Growth of the Imaginal Discs of Eyes of the Wild Type Larvae and Three Mutants—Lobe^c, Glass^2 and Eyeless^2 in *Drosophila melanogaster*," *Zeit. ind. Abst. u. Vererb.*, **70**, 55–72 (1935).

Steinberg, A. G. A Reconsideration of the Mode of Development of the Bar Eye of *Drosophila melanogaster*. *Genetics*, **26**, 325–346 (1941*a*).

Steinberg, A. G., "Studies on Modifier of Bar. A Mutant Which Affects the Expression of the Bar Mutation of *Drosophila melanogaster*," *Genetics*, **26**, 440–451 (1941*b*).

Steinberg, A. G., "The Development of the Wild Type and Bar Eyes of *Drosophila melanogaster*," *Can. Jour. Research*, **D21**, 277–283 (1943).

Steinberg, A. G., and Abramowitz, M., "The Bar 'Locus' and the v^+ Reaction in *Drosophila melanogaster*," *Proc. Nat. Acad. Sci.*, **24**, 107–111 (1938).

MECHANISM OF ENZYMATIC ADAPTATION IN GENETICALLY CONTROLLED YEAST POPULATIONS

BY S. SPIEGELMAN, CARL C. LINDEGREN* AND L. HEDGECOCK

DEPARTMENT OF ZOÖLOGY AND THE HENRY SHAW SCHOOL OF BOTANY, WASHINGTON UNIVERSITY

Communicated December 13, 1943

Introduction.—For some time now it has been recognized that the fermentation of galactose by yeast (see Lippmann[1]) differs greatly from that of other hexoses. Armstrong[2] found that some yeasts had and others had not the power of fermenting galactose. Slator,[3] in a quantitative investigation of the same subject, was able to confirm the statement previously made, that certain yeasts which have the power of fermenting galactose, possess it only after the yeast had been acclimatized by culture in the presence of the sugar.

The nature of this acclimatization has been the subject of some investigation. The experiments of Söhngen and Coolhaas[4] would seem to indicate that the production of galactozymase parallels the formation of new cells. The results of Euler and Nilson[5] also lead to this conclusion. More recently, however, Stephenson and Yudkin[6] concluded from their experiments that the production of galactozymase in a yeast culture need not involve the formation of new cells. They interpret the phenomenon as a direct cytoplasmic reaction to the presence of the substrate (galactose) leading to the formation of a new enzyme. These authors used CO_2 evolution in a period when the population remained constant according to the total and viable counts.

It is clearly of some importance for the further elucidation of this acclimative or adaptive phenomenon to know whether or not new cell formation is involved in the appearance of a new physiological property in a culture. The data presented in this paper indicate that the conflict between the results and conclusions of Stephenson and Yudkin and those of previous workers cited, is more apparent than real. Both types of phenomena ("acclimatization" with and without the formation of new cells) can be observed depending upon the genetic background of the yeast strain being used.

Difficulties in Interpretation.—It must be recognized that microörganisms are handled in such large numbers that when a comparative biochemical investigation of microörganismic physiology is being made, over-all populational characteristics, in contradistinction to the properties of individuals, are being studied. The mechanisms available to an individual for adapting itself to an environmental change are limited by its genome and the physiological flexibility permitted by its particular degree of specialization. When, however, the adaptive ability of a population is being considered, there must be added to the physiological pliability of its members the genetic plasticity of the group in terms of the numbers and kinds of mutants it is capable of producing.

This composite nature of populational adaptability makes the interpretation of observed physiological changes in bacterial or yeast cultures difficult. In any particular case the same end-result can clearly be obtained by any one of the following: (1) the natural selection of existent mutants with the desired physiological characteristic from a heterogeneous population; (2) the presence of the substrate induces a "new"[7] enzymatic system in all the members of a homogeneous population, resulting in a gradual increase in the measured enzyme activity of the population; (3) a combination of natural selection and the action of mechanism (2) on those selected.

In most cases where so-called "training" has been observed to occur in microbic cultures, wherever a decision is at all possible, the most likely explanation seems to be that a natural selection of a variant type has taken place. Relatively few investigations (see Dubos,[8] Yudkin[9] and Karström[10] for recent reviews) can lend much support to the hypothesis of a direct substrate-cytoplasmic interaction.

Thus far, those who have drawn the conclusion that mechanism (2) is operating have depended for the most part on demonstrating the existence of a constant total and viable count during the period when the physiological change in the population was being noted. A constant viable count alone would clearly have little critical value in deciding the question at issue. During the stationary phase of population growth new individuals must make their appearance to take the place of those that have died and natural selection could work in this period as well as in any other of the growth cycle.

However, the use of the constant *total* count involves some difficulties which may briefly be summarized as follows:

1. The failure to find "adaptation" in the complete absence of cell division would not necessarily be conclusive evidence that none exists. It is conceivable that in cultures where this "ideal" had been reached the physiological state of the cells was such that the ability to synthesize new enzymatic systems had been lost along with the ability to divide.

2. This method has the further disadvantage that it cannot distinguish between the following two situations:

(*a*) All the members of a phenotypically homogeneous population react to some component in the environment in an "adaptive" manner.

(*b*) Starting with a phenotypically heterogeneous population only those react which, due to a proper genetic constitution, already possess an adequate though latent enzymatic system. Under this condition providing the genetic variants formed a reasonably high proportion of the population (10% and up) an increased enzymatic activity with time would be observed without any concomitant cell division.

It is evident from the above discussion that the crucial point at issue in questions of this nature is the phenotypic homogeneity or heterogeneity of the starting population. And, unless in a particular case a decision is arrived at on this question, no precise or certain conclusions can be drawn, even under conditions of a constant total count, concerning the mechanism involved in accomplishing an observed change in the physiological characteristics of a population.

Since limitations of size, and sensitivity of available methods, prevent a direct examination of the individuals in a microörganismic population as a means of deciding the question of phenotypic homogeneity, recourse has been made to another method.

A cell can in general be characterized by the type of colony it produces when grown on an agar surface under standard conditions. Thus if a representative sample of a population is plated out and more than one colonial type is noted, it may be logically concluded that more than one type of individual is represented in the population. This method of examination can, under controlled conditions, yield quantitative data of a high degree of reproducibility and has been used to examine the kinetics of the appearance of morphological variants in bacterial cultures (see Shapiro, Spiegelman and Koster[11]).

Methods and Material.—(*A*) *Yeast Strains:* Two strains of *Saccharomyces cerevisiae* isolated in this laboratory and known as Db23B and LK2-G12 were used in the following experiments. Both strains could acquire the ability to ferment galactose if grown in its presence. Strain Db23B originated from a single ascospore and its population contains principally haploid cells and is therefore characteristically unstable. Strain LK2G12 originated from an intact 4-spored ascus in which copulation was observed to occur pairwise. Consequently it is known to be a diploid and in contrast to the Db23B is characteristically stable. For further details concerning yeast breeding and strain isolations and the implications of haploidy and diploidy for genetic stability consult Lindegren and Lindegren.[12]

(*B*) *Media:* The basic medium contains 2 g. of KH_2PO_4, 5 g. of peptone and 2.5 g. of dried dead yeast per liter. To this was added the desired carbohydrate. Agar plates were made by adding the requisite amount of agar to the above basic medium. It was found necessary to filter the agar medium to facilitate observation of the colonies.

(*C*) *Carbohydrates:* As sources of carbohydrate, reagent grade chemicals were used. Difco's purified galactose was further treated according to a method described by Stephenson and Yudkin[6] to remove any contaminating fermentable sugars.

(*D*) *Cell Count:* Cell counts were made by means of an haemocytometer.

(*E*) *Test Plates:* Since we are here concerned with whether or not the organisms can produce CO_2 from galactose, some method had to be devised which would permit the detection of this characteristic. The procedure adopted was to plate a sample of a population on an agar surface, containing 4% galactose purified as described above, and then cover it with another layer of galactose-containing agar. The production of gas cracks the agar right above the colony and thus gas producers can be distinguished easily from those which cannot evolve gas from this sugar. Figure 1 represents a typical field containing both positives and negatives.

To examine the type composition of a population growing on an ordinary glucose plate, well-isolated colonies incubated for 72 hrs. at 26°C. were used. These were removed from the agar surface and suspended in chilled galactose broth and then diluted to contain about 5000 cells per cc. as determined by direct count with an haemocytometer. Of this suspension, 0.2 cc. was then placed on a 4% galactose-4% agar surface from which all excess fluid had been allowed to drain. A sterile bent rod was then rotated over the surface to get even distribution. A 5% agar medium containing 4% galactose cooled to 39° was poured over the inoculated surface. These test plates were incubated at 28°C. for at least 48 hrs. before count-

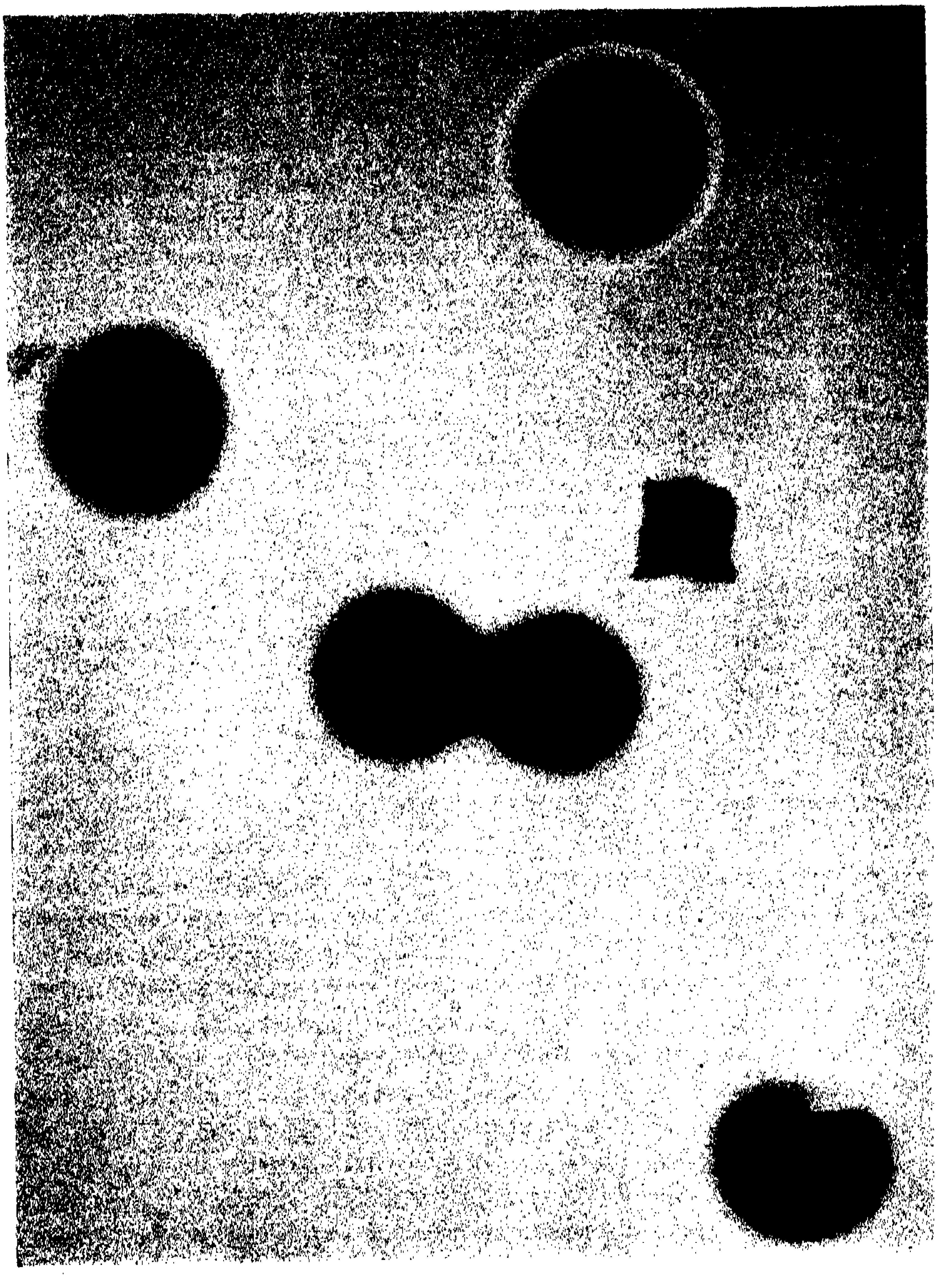


FIGURE 1

Field of a test plate in which colonies grow embedded between two layers of agar. Positives (gas producers) are easily distinguished by the typical star-shaped crack in the agar of its immediate neighborhood.

ing. Five such test plates were prepared from each suspension. Counting was done under the low power of a binocular dissecting microscope.

Experimental Results.—(*a*) *Data obtained from strain Db23B:* Table 1 records results of the galactose test-plate examination of the composition of 10 colonies of Db23B grown on glucose agar. It is clear that two types of individuals, at least as far as gas production from galactose is concerned, are initially present. The percentage of the population, grown on glucose, which is able to ferment galactose, varies from 2.0 to 15.0%.

TABLE 1

COMBINED TABLE SHOWING RESULTS OF EXPERIMENTS DETERMINING THE PERCENTAGE COMPOSITION OF POPULATIONS GROWN ON GLUCOSE WITH RESPECT TO THE ABILITY OF INDIVIDUALS TO PRODUCE CO_2 FROM GALACTOSE. EACH EXPERIMENT REPRESENTS THE COMBINED RESULTS OF FIVE TEST PLATES

	EXPERIMENT NO. 1	TOTAL NUMBER OF COLONIES COUNTED	AVERAGE % OF POSITIVES OBSERVED
Strain Db23B	1	662	7
	2	626	9
	3	622	5
	4	625	15
	5	655	5
	6	599	7
	7	560	4
	8	685	2
	9	530	5
	10	577	13
Strain LK2G12	1	526	100
	2	480	100
	3	692	100
	4	510	100
	5	525	100

Disregarding the quantitative aspect of the problem for the moment, the important point is established that normal populations of Db23B growing on glucose contain two types of individuals which can be classified on the basis of the behavior of their colonies when grown on galactose agar test plates.

(*b*) *Data obtained with LK2G12:* Table 1 also records some representative results obtained from the examination of LK2G12 colonies grown on glucose agar. It is evident that here we are dealing with a population homogeneous with respect to the observed characteristic.

(*c*) *Validity of method:* It might be argued that whether a particular colony exhibited gas production by a crack in the agar might depend on local characteristics of the agar in which it found itself.

Two types of experiments were undertaken to examine this question. In the first type 5% glucose rather than galactose test plates were prepared and seeded in exactly the same fashion with Db23B which was known to ferment glucose with great ease, as determined from manometric experiments on this strain in an atmosphere of nitrogen.

If the appearance of a crack immediately above a colony did depend on local imperfections, then the same general picture should be obtained in these glucose plates as was noted in the galactose test plates. In all, 11 test plates were so prepared and 2592 colonies examined. Of these only one did not show a strong positive reaction as exhibited by the typical star-shaped crack. This latter colony when isolated and examined manometrically appeared to have the non-fermentative type of glucose oxidation.

TABLE 2

EXPERIMENTS TESTING VALIDITY OF DETECTING METHOD. PERCENTAGES ARE CALCULATED TO THE NEAREST INTEGER. EACH FIGURE REPRESENTS THE AVERAGE OF FIVE TEST PLATES

EXPERIMENT NO.	% POSITIVES		VOLUME RATIO OF SUSP. 1:SUSP. 2 USED IN MIXTURES
	SUSPENSION 1	SUSPENSION 2	
1	20	48	1:1
2	4	98	1:1
3	10	67	1:1
4	10	67	1:2

EXPERIMENT NO.	COMPOSITION OF MIXTURES		DIFFERENCE CALCULATED − OBSERVED
	OBSERVED % POSITIVES	CALCULATED % POSITIVES	
1	31	34	3
2	48	51	3
3	38	38	0
4	50	48	−2

As a final test, the following experiment was performed: it was found, as was to be expected, that the per cent of gas-producing colonies obtained in a test plate originating from a positive colony on galactose was much higher than either from a negative one on galactose or a normal one growing on glucose. The percentage of positives obtained from such a colony varied from 75 to 99% depending on the length of incubation, a general rise with time being noted. Suspensions of equal density as determined by cell count were made of a normal colony from glucose and a positive one from a galactose plate. Five test plates were made from samples of each suspension to determine the percentage composition of the population. They were then mixed in 1:1 and 1:2 proportions and five galactose test plates made from samples of the mixtures. If the appearance of the cracks is truly a characteristic of the colony and not one of the locality in which it finds itself, it should be possible to calculate the percentages of positives to be found in the test plates of the mixtures from the percentage composition and proportions used of the two suspensions that went to make it up. The results of these experiments are recorded in table 2.

It is seen that in the first three experiments, within an average difference of 1.5 from the calculated value, the observed per cent positives in the test plates of the mixtures are the arithmetic means of the percentages of the two suspensions from which they originated. In the fourth experiment in which a 1:2 proportion was used, the deviation from the calculated value is 2%.

(*d*) *Measurements of enzyme activity:* As measured by gas evolution in nitrogen, both strains can equally well acquire the ability to ferment galactose if they are grown in the presence of the sugar. Evidently, then, increase in cell number is sufficient to effect the physiological change in either population. The problem is, however, whether the change will occur without this mechanism.

The results on the test plates indicate that Db23B grown on glucose consists mainly of individuals that cannot ferment galactose. On this basis it would appear probable that such a population would have to depend on cell division to change its over-all character with respect to this property by increasing through selection the proportion of individuals that can ferment the sugar. Thus a population of this nature should not "adapt" in the presence of galactose if cell division were either completely or significantly depressed.

Such experiments were performed on Db23B. To keep multiplication down to a minimum the cells were washed in $m/15$ KH_2PO_4 and resuspended in this solution with purified galactose as the only addition. Under such conditions cell counts remain constant over long periods of time. Anaerobic gas production was followed by the usual manometric methods. So long as the total cell count remained constant no change was observed from the very low rate of (endogenous) CO_2 evolution. As pointed out previously, negative results such as these have little crucial interpretive value. They do, however, become more meaningful when they are combined with the test-plate data, and the fact that this strain (in common with many others cited in the literature) invariably adapts to the substrate when allowed to proliferate in its presence.

The results obtained with LK2G12 under the same conditions are entirely different. With this strain, adaptation occurs quickly and almost explosively as may be noted from figure 2. In this experiment exactly the same experimental conditions were imposed as obtained in the experiments on Db23B. As can be seen, a very sharp change in the physiological characteristics of the population took place at the end of about 3 hrs. During this period no measurable change in the total number of individuals occurred as determined by total counts at the beginning and end of the run. These are plotted on the same graph. We have here a clear cut case of "adaptation" without the formation of new cells. It may be mentioned that this lag period of 3 hrs. is under standard conditions, reproducible,

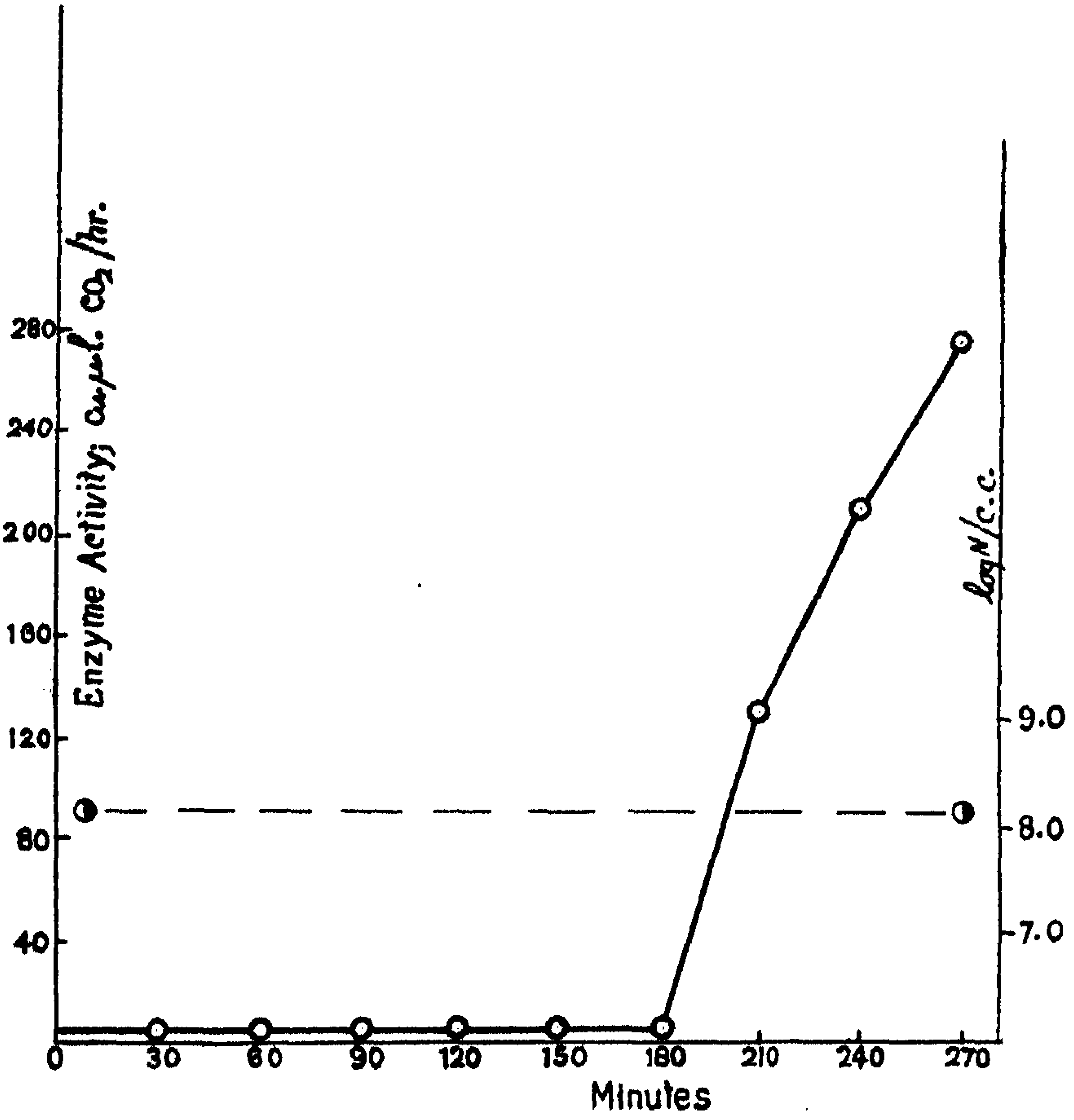


FIGURE 2

Change in physiological character of a stationary LK2G12 population in contact with galactose. ⊙ = Enzyme activity, ◑ = Log of number of cells per cubic centimeter.

and is characteristic for this strain. Other strains examined under the same conditions have longer lag periods extending up to 8 hrs.

Discussion.—These results are not only compatible with the previous test-plate examination of the phenotypic homogeneity of the populations but, in turn, the phenotypic pictures coincide well with what is known about the genetic background of the two strains employed; i.e., Db23B being principally composed of haploid cells would tend to be genetically unstable and consequently would give rise to phenotypically heterogeneous populations so long as vegetative reproduction is maintained.

These experiments offer a reasonable explanation for the apparently contradictory results obtained by various workers on the question of the rôle of cell division in the "acclimatization" of yeast populations to galactose fermentation. Apparently Söhngen and Coolhaas,[4] as well as Euler[5] and his co-workers, were dealing with a culture of the Db23B type whereas Stephenson and Yudkin[6] probably employed a type represented by LK2G12, although with a much longer lag period.

The experiments described were specifically designed to investigate the biological mechanisms underlying populational "adaptations." As such they cannot provide solutions to the many biochemical problems of a more or less fundamental nature which have arisen during the course of the investigation. One important point may, however, be noted here which has as yet received too little attention; namely, that CO_2 evolution need not parallel galactose utilization as a source of energy for growth and maintenance. Of 34 strains of yeast examined in this laboratory, only one showed inability to grow in a medium containing galactose as sole source of carbohydrate, yet 80% of those that could use this sugar failed to ferment it. In particular the utilization of galactose by the "unadapted" culture of Db23B (glucose grown) was examined very carefully by following galactose disappearance during non-fermentative periods.

The existence of these non-fermentative oxidations of carbohydrate has been demonstrated in other forms using other sugars. Trautwein and Wiegend[13] have shown the direct oxidation of maltose by certain molds. Recently Barron and Freidemann[14] have demonstrated the oxidation of glucose and hexose-phosphates by bacteria which do not ferment glucose. As previously noted, a yeast strain possessing this type of glucose oxidation was isolated from one of our experimental glucose test plates.

With this point in mind, it is clear that the present investigation, in common with previous ones on "galactose adaptation," is not concerned with adaptation to galactose utilization *per se*, but with the development of a particular enzyme system leading to a fermentative oxidation of the sugar. The characteristics of the non-fermentative oxidations of galactose will be discussed elsewhere.

Whether it is appropriate to call the lag phenomenon observed with the LK2G12 strain in the presence of galactose an "adaptation," cannot be decided without an opportunity to examine the biochemical data in greater detail. Nor is it here possible to evaluate properly the usefulness of Karström's[10] classification of galactozymase as an "adaptive enzyme" in contradistinction to the glucose fermenting system as a "constitutive" one. Data obtained in this laboratory on the kinetics of the synthesis and destruction of the galactose fermenting system in a genetically stable population, would rather suggest that one was dealing with an enzyme system extremely unstable in the absence of its specific substrate and that what Karström calls a "constitutive" enzyme is simply one that is relatively more stable.

Whatever the case may be, the results reported here would suggest that a proper investigation into the nature of the acclimatization towards CO_2 production from galactose requires some knowledge of the genetic background of the strain being employed as well as an examination of the phenotypic homogeneity of the starting population. Unless this is done, there is the ever-present risk of unknowingly complicating the experiments by studying two distinct phenomena at the same time, one involving the competitive replacement of one phenotype by another and the other the kinetics of the enzyme substrate interaction. The latter problem can be studied safely only in a genetically stable and homogeneous population.

Summary and Conclusions.—1. A method is described involving the use of double-layered agar test plates which permits an examination of the phenotypic homogeneity of yeast populations with respect to the ability of individual members to produce gas from galactose. This method was used to elucidate the mechanism of enzymatic adaptation in a haploid and in a diploid strain of yeast.

2. Adaptation in the haploid was found to occur only in increasing populations.

3. Adaptation in the diploid could occur in the absence of cell division.

4. The necessity for a knowledge of the genetic background and phenotypic composition of initial populations is emphasized, to avoid complicating studies of enzyme synthesis with the kinetics of competitive interaction between phenotypes.

[*] Aided by a grant from Anheuser-Busch, Inc., St. Louis.

[1] Lippmann, E. V., *Die Chemie der Zuckerarten*, Braunschweig, 1904.

[2] Armstrong, E. F., *Proc. Roy. Soc.*, **B76**, 600–605 (1905).

[3] Slator, A., *Jour. Chem. Soc.*, **93**, 217–242 (1908).

[4] Söhngen, N. L., and Coolhaas, C., *Jour. Bact.*, **9**, 131 (1924).

[5] V. Euler, H., and Nilson, R., *Zeit. Physiol. Chem.*, **143**, 89–107 (1925).

[6] Stephenson, M., and Yudkin, J., *Jour. Biochem.*, **30**, 506–514 (1936).

[7] The word "new" is here used advisedly and is not meant to imply that the enzyme system formed was necessarily entirely foreign to the cell. Experimentally, due to the limitations of available techniques, it is impossible to distinguish between this situation and one where the enzyme was actually present but in too small an amount to be measured.

[8] Dubos, R. J., *Bact. Rev.*, **4**, 1–16 (1940).

[9] Yudkin, J., *Biol. Rev. Camb. Phil. Soc.*, **13**, 93–106 (1938).

[10] Karström, J., *Ergeb. Enzymforsch.*, **7**, 350–376 (1937).

[11] Shapiro, A., Spiegelman, S., and Koster, H., *Jour. Genetics*, **34**, 237–245 (1937).

[12] Lindegren, Carl C., and Lindegren, G., *Ann. Mo. Botan. Garden*, **30**, 453–468 (1943).

[13] Trautwein, K., and Wiegand, K., *Biochem. Z.*, **240**, 423 (1931).

[14] Barron, Guzman, E. S., and Friedemann, T. E., *J. Biol. Chem.*, **137**, 593–615 (1941).

ERRATA

In the article, "On Convex Topological Linear Spaces," by GEORGE W. MACKEY, **29**, 315–319 (1943), in lines 22 and 23 the phrase "whose continuous linear functionals" should be replaced by "the pseudo-norm sets of whose continuous pseudo-norms."

In the article, "Natural Logarithms of Small Prime Numbers," by H. S. UHLER, **29**, 319–325 (1943), in line 2 up from the bottom of page 323 in the value of ln 43 the ninth pentad "12917" should be changed to read "12971."

PROCEEDINGS
OF THE
NATIONAL ACADEMY OF SCIENCES

Volume 30 February 15, 1944 Number 2

Copyright 1944 by the National Academy of Sciences

RELATIVE NUMBER OF NON-INVARIANT OPERATORS IN A GROUP

By G. A. Miller

Department of Mathematics, University of Illinois

Communicated December 30, 1943

Since the number of the invariant operators of a given group G, of order g is equal to the order of the central c of G, the number of the non-invariant operators of G is $g - c$. In every non-abelian group the central quotient group is known to be non-cyclic and hence its order is a composite divisor of the order of G. In particular, when the order of G is the product of two distinct prime numbers and G is non-abelian the number of the non-invariant operators of G is $g - 1$, and when the order of the non-abelian group G is p^3, p being a prime number, then the number of the non-invariant operators of G is $p^3 - p = p(p^2 - 1)$. A non-abelian group of order g can therefore not contain less non-invariant operators than g minus g divided by the two smallest positive integers which divide the order of g.

From the preceding paragraph it results that a non-abelian group of order g cannot have less than $3g/4$ non-invariant operators and that it must involve a larger number of such operators when g is not divisible by 4. In fact, it must involve a larger number of such operators unless g is divisible by 8 as we proceed to prove. Suppose that G contains exactly $3g/4$ non-invariant operators and hence its central quotient group is the four group. Its central will then give rise to three co-sets corresponding to the three operators of order 2 in the four group. If s_1 is an operator from one of these co-sets and s_2 is an operator from another of these three co-sets then s_1 and s_2 are necessarily non-commutative sines; otherwise they would appear in the central of G.

A commutator which arises from s_1 and s_2 must be in the central of G and it must be of order 2 since the squares of all of the operators of G must appear in the central of G. It therefore results that the central of G is of even order and hence the order of G is divisible by 8. It is obvious that each of the two non-abelian groups of order 8 satisfies the condition that the number of the non-invariant operators contained therein is $3g/4$, where

g is the order of the group. Moreover the direct product of one of these groups of order 8 and an arbitrary abelian group evidently satisfies the same condition. Hence there results the following theorem: *A necessary and sufficient condition that there is at least one group of order g in which the number of the non-invariant operators is exactly $3g/4$ is that g is divisible by 8 and there is no non-abelian group of order g which contains a smaller number of non-invariant operators.* This theorem can clearly be extended.

When the number of the non-invariant operators of G is exactly $3g/4$ all the operators of odd order contained in G must appear in its central. Hence such a G is the direct product of its Sylow subgroups and all of its Sylow subgroups of odd order are abelian while its Sylow subgroup whose order is a power of 2 has a commutator subgroup of order 2 as was noted above. The determination of the groups which involve exactly $3g/4$ non-invariant operators is therefore reduced to the determination of such groups having an order which is a power of 2. In particular, if a group whose order is not divisible by 16 has the property that it contains exactly $3g/4$ non-invariant operators it is the direct product of one of the two non-abelian groups of order 8 and an arbitrary abelian group of odd order, and all such direct products have the property that the number of the non-invariant operators contained in each of them is $3g/4$, where g is the order of the group.

If G contains exactly $3g/4$ non-invariant operators then it contains three abelian subgroups of order 2^m, $m > 1$, which have 2^{m-1} operators in common and these common operators are contained in the central of G while the Sylow subgroup whose order is a power of 2 is of order 2^{m+1}. The commutator subgroup of order 2 of G appears in the central of G but it is not necessarily generated by an operator of higher order in G as results from a non-abelian group of order 16 in which two of the three abelian subgroups of order 8 are of type 2, 1 and the central is the four group. Two of the three operators of order 2 in this four group are squares of operators of order 4 contained therein while the third is a commutator of the group. When the number of the non-invariant operators of a group is $3g/4$ the smallest number of these operators is 6 since it must be an even number and the two non-abelian groups of order 8 are the only groups which have separately exactly 6 non-invariant operators.

Every group of order g which has the property that exactly $3g/4$ of its operators are non-invariant contains a multiple of 6 non-invariant operators and it is possible to construct such a group in which the total number of non-invariant operators is an arbitrary multiple of 6. This results directly from the fact that in the direct product of a group which involves exactly 6 such operators and an abelian group of arbitrary order the number of the non-invariant operators is equal to 6 times this arbitrary order. In particular, the groups which contain exactly 12 non-invariant operators are of order 16 and have a central of order 4. There is one and only one such group which contains the cyclic subgroup of order 8. The remaining operators of this group transform each operator of the cyclic group into its fifth power. There is obviously another such group in which the central is the cyclic group of order 4 and there are four such groups in which the central is the non-cyclic group of order.

There is only a finite number of different non-abelian groups which separately have the property that each of them contains the same number of non-invariant operators. This number of groups may be zero. The fact that there is only a finite number of such different groups is a direct consequence of the theorem that there is an upper limit for the order of such groups since this order cannot be as large as twice the number of the non-invariant operators contained therein. On the contary, the number of the different groups which contain separately the same given number of non-invariant subgroups is not necessarily limited by this number of the non-invariant subgroups contained in these separate groups. In fact, it is known that there is an infinite system of groups such that each of them contains two and only two non-invariant subgroups. (Cf. these *Proceedings*, **29**, 105 (1943).) There is, however, no group which contains exactly two non-invariant operators as will be more fully explained.

If the number of the non-invariant operators of G exceeds $3g/4$ it is at least $5g/6$ because the order of the central quotient group is then at least 6 and if this order is 6 the central quotient group is the non-cyclic group of this order. The direct product of this group and an arbitrary abelian group evidently furnishes an infinite system of groups such that the number of the non-invariant operators contained in each of these groups is $5g/6$, g being the order of the group. The number of the non-invariant operators in each of these groups is a multiple of 5 and there is at least one group in which this number is an arbitrary multiple of 5. The symmetric group of order 6 is the only one of these groups which contains exactly 5 non-invariant operators and this group is characterized by the fact that it contains exactly 5 non-invariant operators.

A necessary and sufficient condition that there exists a group in which the number of non-invariant operators is exactly a given prime number p is that there exists a group of order $p + 1$ which contains no invariant operator besides the identity. This theorem results from the fact that this prime number is equal to the order of the central quotient group diminished by unity and that this central quotient group cannot involve any invariant operator besides the identity. It therefore results that there exists no group in which the number of the non-invariant operators is exactly 3, 7 or 31, but that there exists at least one non-abelian group in which the number of the non-invariant operators is any one of the prime numbers in the following set: 5, 11, 13, 17, 19, 23, 29. It would be a very simple matter to extend these sets.

It is now easy to prove the following theorem: *Every group in which the total number of non-invariant operators is a prime number contains no invariant operator besides the identity.* In fact, if the total number of the non-invariant operators of the group is a prime number the order of the central quotient group diminished by unity must be equal to this prime number. The central quotient group must be simply isomorphic with the group since the number of the non-invariant operators of a group is the product of the order of the central, and the order of the central quotient group diminished by unity. The converse of this theorem is obviously not necessarily true since it is easy to find groups which contain no invariant operator besides the identity but in which the number of the non-invariant operators is not a prime number. For instance, the dihedral group of order 10 has this property.

There is one and only one group in which the total number of the non-invariant operators is one of the three prime numbers 5, 11, 13 but there are two groups in which the number of the non-invariant operators is exactly 17. One of these is the dihedral group of order 18 and the other is the generalized dihedral group of this order. When the number of the non-invariant operators contained in G is exactly 23 the symmetric group of order 24 is the only one of the 15 groups of this order which contains no invariant operator besides the identity and hence this group is characterized by the fact that it contains exactly 23 non-invariant operators. This characterization of the symmetric group of degree 4 may therefore be added to the numerous known definitions of this well-known group.

NOTE ON THE FUNCTION $F(a, b;\ c - n;\ z)$

By H. Bateman*

Norman Bridge Laboratory of Physics, California Institute of Technology

Communicated January 8, 1944

The generating function

$$(1 - t)^{b-1}(1 - z + zt)^{-a} = \sum_{n=0}^{\infty} (t^n/n!)(1 - b, n)F(a, b;\ b - n;\ z) \quad |t| < 1,\ |zt| < |1 - z|$$

may be used to find an estimate of $F(a, b;\ b - n;\ z)$ for large positive values of n. When the point $t = 1 - 1/z$ lies outside the circle $t = 1$ the singularity $t = 1$ may be used to find an estimate by the method of Darboux[1] and the result is

$$F(a, b;\ b - n;\ z) \sim 1.$$

When $c - b$ is an integer the foregoing result indicates that $F(a, b;\ c - n;\ z) \sim 1$ under the same restriction on z. This is the result used in previous papers[2] but the case in which the point $t = 1 - 1/z$ lies within the unit circle or on its circumference was not considered. These cases can also be treated by the method of Darboux. It may be noted, however, that when $1 - 1/z$ lies inside the unit circle the point $1 - 1/(1 - z)$ lies outside the circle. Use may then be made of the formula of Barnes[3]

$$F(a, b;\ c - n;\ z) = AF(a, b;\ 1 + a + b - c + n;\ 1 - z) + \\ Bz^{1 - c + n}(1 - z)^{c - a - b - n}F(1 - a, 1 - b;\ 1 - a - b + c - n;\ 1 - z)$$

where

$$A = \Gamma(c - n - a - b)\Gamma(c - n)/\Gamma(c - n - a)\Gamma(c - n - b)$$
$$B = \Gamma(a + b - c + n)\Gamma(c - n)/\Gamma(a)\Gamma(b).$$

It should be noticed that $(1 - a - b + c) - (1 - a)$ is an integer when $c - b$ is an integer so the previous result may be applied. It is probable, however, that the result holds without this restriction. The first term is $0(1)$ when n is large and so in this case we have the estimate

$$F(a, b;\ c - n;\ z) \sim n^{a + b - 1}(1 - z)^{1 - a - b}\left(\frac{z}{1 - z}\right)^{1 - c + n}\pi \div \\ [\Gamma(a)\Gamma(b) \sin \pi(c - n)].$$

On this account restrictions are needed for the deduction $\lim\limits_{n \to \infty} F(-m, -r;\ -n - r;\ x) = 1$. When the point $1 - 1/z$ lies on the unit circle there are two possible singularities to be taken into consideration in the method of Darboux and so unity should be added to the foregoing expression. If, however, $R(a + b) < 1$ unity is negligible in comparison with the term just mentioned and if $R(a + b) > 1$ unity is the dominant term. The case in which $R(a + b) = 1$ is exceptional because then both terms must be taken into consideration.

It is noteworthy that in the case $z = {}^1/_2$ there are two singularities on the unit circle. For the function of Mittag-Leffler

$$\begin{aligned} g_n(z) &= 2zF1 - n, 1 - z;\ 2;\ 2) \\ &= 2^{-z}(z/, n)F(z + 1, z;\ 1 + z - n;\ {}^1/_2) \end{aligned}$$

the condition $R(a + b) < 1$ is satisfied when $R(z) < 0$ and in this case an estimate is

$$g_n(z) \sim 2^{-z}(z/, n).$$

When $R(z) < 0$ use may be made of the formula

$$g_n(-z) = (-)^n g_n(z),$$

so in this case

$$g_n(z) \sim (-)^n 2^z(-z/, n) = 2^z(z, n)/n.$$

For the function

$$\begin{aligned} g_n(z, r) &= (-)^n(r/, n)F(-n, -z - r, -r; 2) \\ &= (z + r/, n)2^{-z}F(z, r + z + 1; r + z - n + 1; {}^1/_2) \end{aligned}$$

the condition $R(a + b) < 1$ is satisfied when $R(2z + r) < 0$ and so the estimate

$$g_n(z, r) \sim (z + r/, n)2^{-z}$$

holds under this condition. The equation

$$g_n(-z - r, r) = (-)^n g_n(z, r)$$

indicates that when $R(2z + r) > 0$ the estimate is

$$g_n(z, r) \sim (-)^n 2^{z + r}(-z/, n) = 2^{z + r}(z, n)/n!.$$

When $R(2z + r) = 0$ the proper estimate is

$$g_n(z, r) = (z + r/, n)2^{-z} + 2^{z + r}(z, n)/n!$$

and in the case of the polynomial of Mittag-Leffler, when $R(z) = 0$

$$g_n(z) \sim 2^{-z}(z/, n) + 2^z(z, n)/n!$$

---

* For list of Errata to previous articles, see page 44, infra.

[1] Darboux, G., *Jour. de Math.*, (3) **6**, 1–56, 377–416 (1878).

[2] Bateman, H., *Proc. Nat. Acad. Sci.*, **26**, 491–496 (1940); **28**, 371–374 (1942).

[3] Barnes, E. W., *Proc. London Math. Soc.*, (2) **6**, 157 (1908).

INDOLE AND SERINE IN THE BIOSYNTHESIS AND BREAKDOWN OF TRYPTOPHANE*

By E. L. Tatum and David Bonner

Department of Biology, Stanford University

Communicated January 31, 1944

Indole and anthranilic acid have been suggested as intermediates in the biosynthesis of tryptophane by certain bacteria.[1, 2] Among the mutant strains of *Neurospora*, produced as described elsewhere,[3] two types of tryptophane deficient strains have been found. One strain will grow in the presence of indole, and the other in the presence of either indole or anthranilic acid. With the use of these mutants it has been shown that both of these compounds are intermediates in the synthesis of tryptophane by *Neurospora* and that this synthesis has been blocked before anthranilic acid in one mutant and between anthranilic acid and indole in the other (10575).[4] This paper deals with the nature of the reaction connecting indole with tryptophane. Certain of the results with *Neurospora* have been included in a preliminary note.[5]

Synthesis of Tryptophane by Neurospora.—Tests of a number of possible intermediates between indole and tryptophane, including skatole, indoleacetic acid, indolepropionic acid, l-kynurenine, tryptamine, indole-lactic acid, and indolepyruvic acid (sterilized by filtration) showed that none was active in promoting growth of the tryptophane deficient strain 10575. These inactive compounds cannot be converted to either indole or tryptophane, and presumably are not intermediates in the synthesis of tryptophane by *Neurospora*. Since these substances include the theoretically possible intermediates in a stepwise synthesis of tryptophane from indole, it seemed, therefore, that this synthesis might involve a single reaction.

In an attempt to determine the nature of this step in the synthesis of tryptophane from indole, a number of experiments were carried out both with actively growing cultures of *Neurospora* and with mycelium in non-nutrient media. Indole was determined by treating samples of culture media with *p*-dimethylaminobenzaldehyde essentially as described by Happold and Hoyle,[6] and measuring the color developed with a photoelectric colorimeter. The disappearance of indole and the effects of various substances on this disappearance were followed by this method. The experimental results may be summarized as follows. Indole disappears from the culture solutions rather slowly. The addition of serine causes the indole to disappear much more rapidly. The action of serine is specific: α or β alanine, pyruvic acid, glyceraldehyde, phosphoglyceric acid, glycolic acid, threonine, Cori-ester and sucrose are without effect. *l*(−)-serine[7] has twice the activity of *dl*-serine. When serine is limiting, both the rate of disappearance of indole and the amount of indole used are functions of the serine concentration as shown by representative data presented in figure 1. Results essentially similar to these have been obtained with the normal wild-type strain of *Neuorspora*. It is therefore clear that *l*(−)serine is involved in the utilization of indole by *Neurospora*.

Since *Neurospora* grown on indole has been found to contain tryptophane, it seemed possible that this synthesis might be taking place under the experimental conditions used in this investigation. Colorimetric tests for tryptophane in the solutions after disappearance of indole supported this interpretation. Further experiments were therefore undertaken. Cultures of the wild-type and mutant strains were grown with constant shaking. After 3 days mycelia representing about 1 gm. dry material were washed and incubated with 50 mg. indole and 500 mg. serine in 1000 ml. of distilled water. After 48 hours at 25°C. with constant shaking the indole had disappeared. Tryptophane was then determined in the medium essentially as described by Woods,[8] but with a photoelectric colorimeter. The results indicated that tryptophane was present in yields of 30 to 50 per cent of theory, while the indole content of control cultures without serine had decreased very little, and no tryptophane could be detected. The results of biological assay for tryptophane on the indole-serine solutions using strain 10575 agreed well with the results of the colorimeter assay. This shows that the product is predominantly *l* (–)-tryptophane, since the *d*(+) form is inactive for *Neurospora* strain 10575. In other control cultures to which *l* (–)tryptophane and serine had been added, about 40 per cent of the tryptophane remained after 48 hours' incubation. It therefore seems evident that all of the indole was converted into *l* (–) tryptophane, but that 50 to 70 per cent of the tryptophane was broken down during the incubation period. One of the products of this decomposition of tryptophane was found to be indole-acetic acid.[9] Indole-

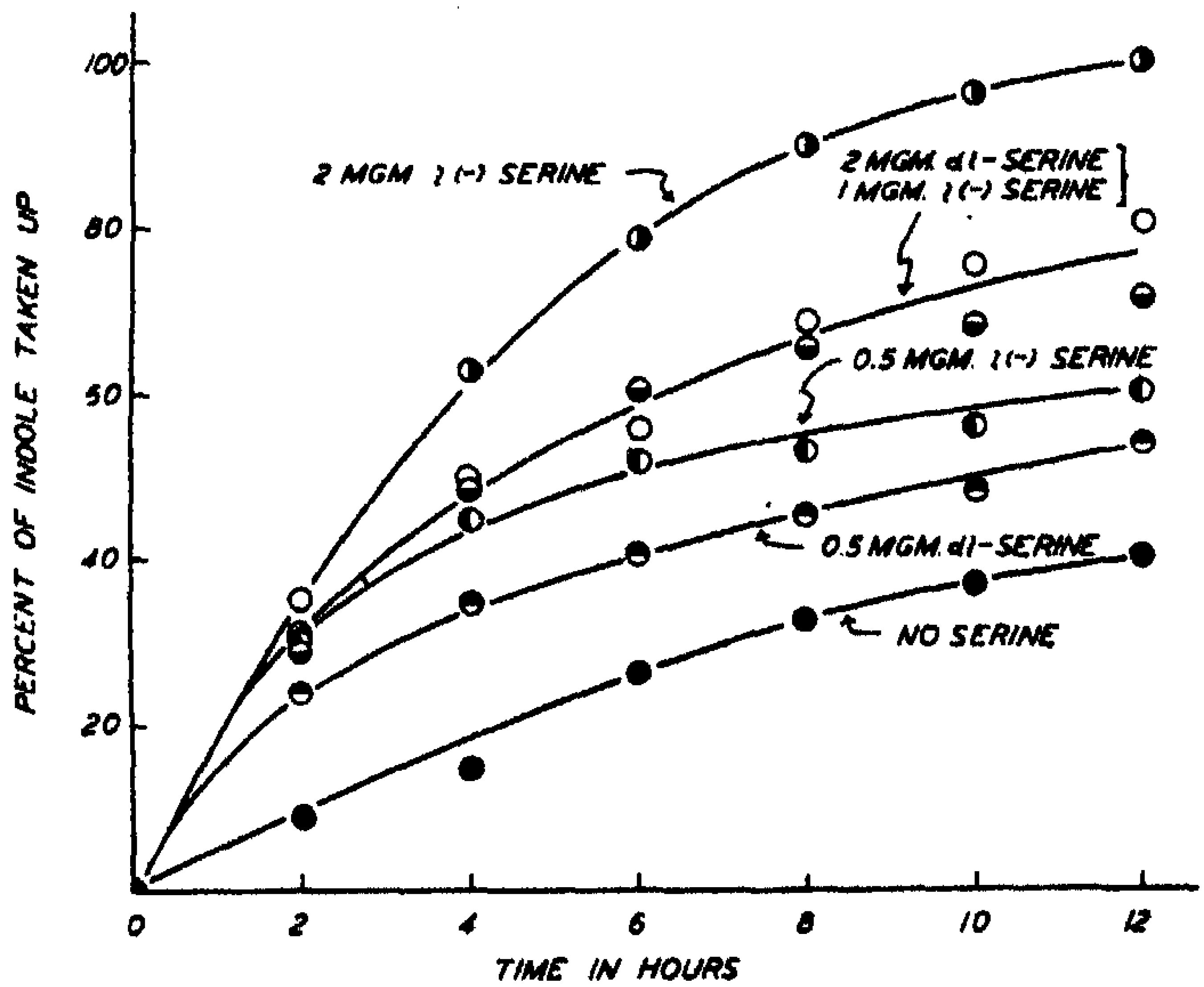


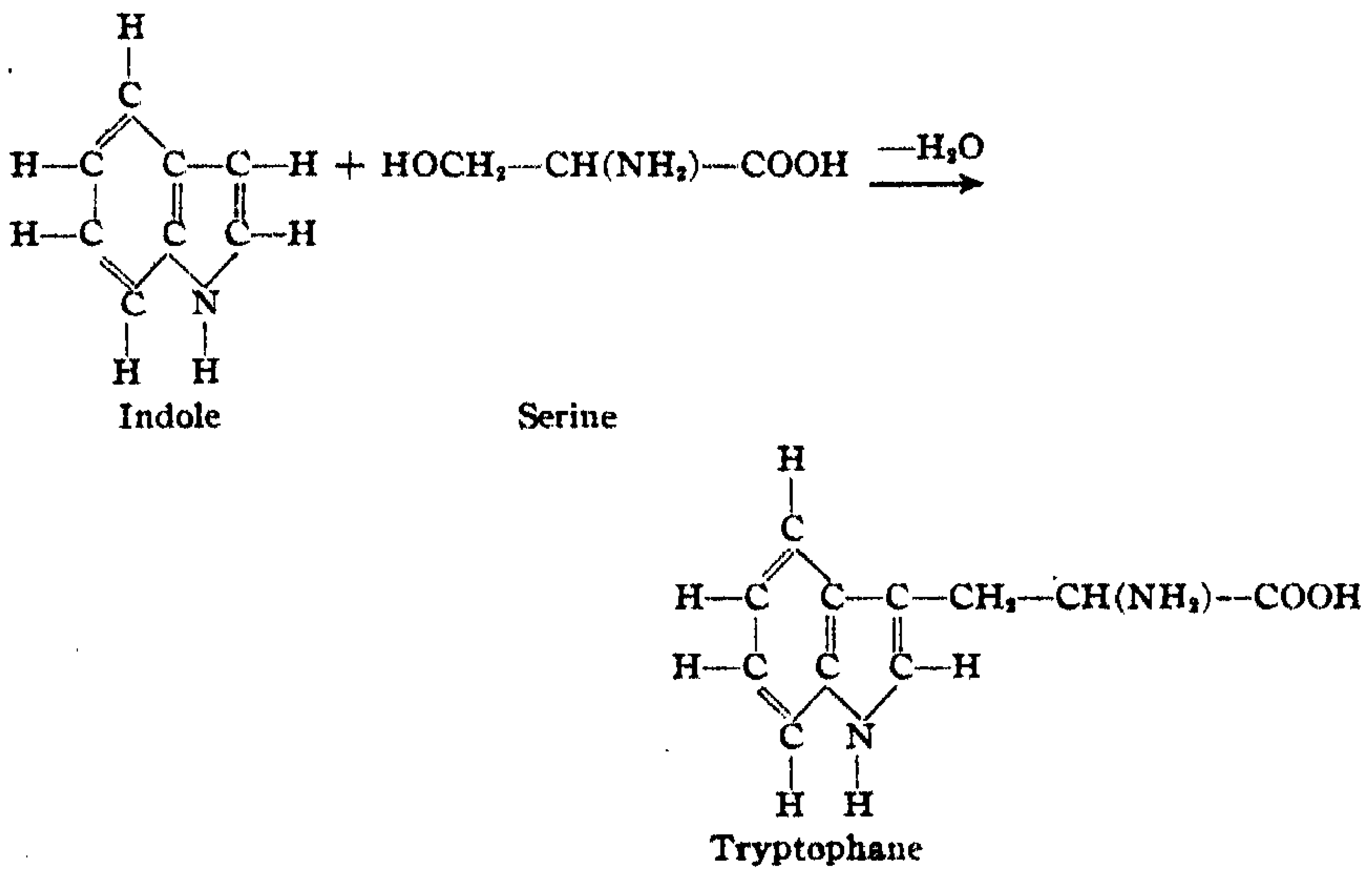
FIGURE 1

The effect of serine on the uptake of indole by *Neurospora* mycelium. Each 125-ml. culture flask contained 1 mg. indole and *ca.* 50 mg. of 3-day-old washed mycelium of strain 10575 in 50 ml. distilled water. The flasks were stoppered and incubated at 25°C. with constant shaking.

acetic acid is produced and excreted by *Neurospora* under the conditions of the experiments only in the presence of tryptophane or of a mixture of indole and serine. The postulation of the synthesis of tryptophane in a reaction involving indole and serine is further supported by the observation that in growth experiments *Neurospora* will tolerate higher concentrations of indole in the presence of added serine. Tryptophane is much less toxic than indole.

Although the evidence strongly indicated that indole reacts with serine to give tryptophane, it is possible that neither the colorimetric nor the biological assays were measuring tryptophane specifically. Final proof was obtained by the actual isolation of tryptophane from the incubation mixture.

A solution which by colorimetric and biological assays contained about 40 mg. of tryptophane was concentrated in vacuum, precipitated with $HgSO_4$, the precipitate freed of mercury by treatment with $Ba(OH)_2$ and H_2S, and the H_2S removed in vacuum. The solution was adjusted to pH 6 with $Ba(OH)_2$ and H_2SO_4, and concentrated under reduced pressure. It proved difficult to crystallize the product from this solution and therefore it was treated with acetic anhydride as described by Berg, Rose and Marvel.[10] This procedure leads to the racemization of tryptophane with the production of acetyl *dl*-tryptophane.[11] The acetylated product was extracted from acid solution with ether and the ether removed. Crystallization from this solution was difficult until it was found that very small amounts of indole inhibited the crystallization of acetyl-tryptophane. After treating the solution with a little Norite the product crystallized easily from water. Approximately 15 mg. of crystalline material were obtained. After purification by recrystallization this product gave the correct melting point for acetyl-*dl*-tryptophane (204–205°C.) and showed no depression when mixed with authentic acetyl-*dl*-tryptophane. Analysis[12] gave the following results: Calculated for $C_{13}H_{16}N_2O_3$: 63.10% C; 6.12% H; 11.34% N. Found: 63.74% C; 5.94% H; 11.73% N. The isolated material therefore was acetyl-*dl*-tryptophane, and the product of the reaction between indole and serine was tryptophane. Since the colorimetric and biological assay values agreed this presumably was the biologically active *l*(−)isomer. While it is possible that serine does not actually enter into the reaction but merely permits the accumulation of tryptophane after its formation from indole by some other reaction, this does not seem likely. The structural relationship of serine to the side chain of tryptophane, the specificity of serine in the reaction, and the failure to detect tryptophane in an incubation mixture with indole or serine alone make it probable that in *Neurospora* indole reacts with *l*(−)-serine to give *l*(−)tryptophane. One possibility is that an intermediate such as α-aminoacrylic acid[13] is involved. Since only *l*(−)serine is active serine could react directly with indole with the elimination of water as follows:



Since all possible intermediates tested failed to support growth of mutant strain 10575, it seems probable that the condensation of indole and serine is involved in the normal synthesis of tryptophane by *Neurospora.*

Production of Indole from Tryptophane by Escherichia Coli.—The formation of indole from tryptophane by bacteria has been studied by a number of investigators.[6, 8, 14, 15] In this reaction no suspected intermediates tested give rise to indole directly. Indolelactic and indolepyruvic acids according to Woods may give indole, but must first be converted to tryptophane.[8] *E. coli* resembles *Neurospora* in that skatole, indoleacetic acid, indolepropionic acid, 1-kynurenine and tryptamine cannot be converted into tryptophane.

It seemed possible that the bacterial production of indole from tryptophane might be a simple reversal of the synthetic reaction carried on by *Neurospora.* If this were true it should be possible to modify the production of indole through the addition of serine.

Sterile-washed suspensions of *E. coli* (K-12)[16] were prepared as described by Woods.[8] The cell suspensions were incubated with tryptophane in flasks on a shaking machine and indole determined colorimetrically at intervals on aliquot samples. Corrections were made for the turbidity due to the cells.

The results of a typical experiment are shown in figure 2. The addition of *dl*-serine greatly slows down the production of indole. As with *Neuro-*

spora, the effect of serine is specific, and is a function of the serine concentration. In our experiments glucose and α-alanine have no influence on the reaction. Evans, *et al.*, found that of twelve amino acids (not including serine) tested in growth experiments, the production of indole from tryptophane was effected only by phenylalanine and tyrosine, which two amino acids, however, had no influence on the production of indole by active cell suspensions.[14] We have also performed a few experiments in which indole alone, and indole with serine were incubated with coli suspensions. The results showed clearly that serine has an effect on the disappearance of indole similar to that in *Neurospora*. Colorimetric tests for tryptophane after 24 hours' incubation indicated that some tryptophane had been synthesized from indole in the presence of serine.

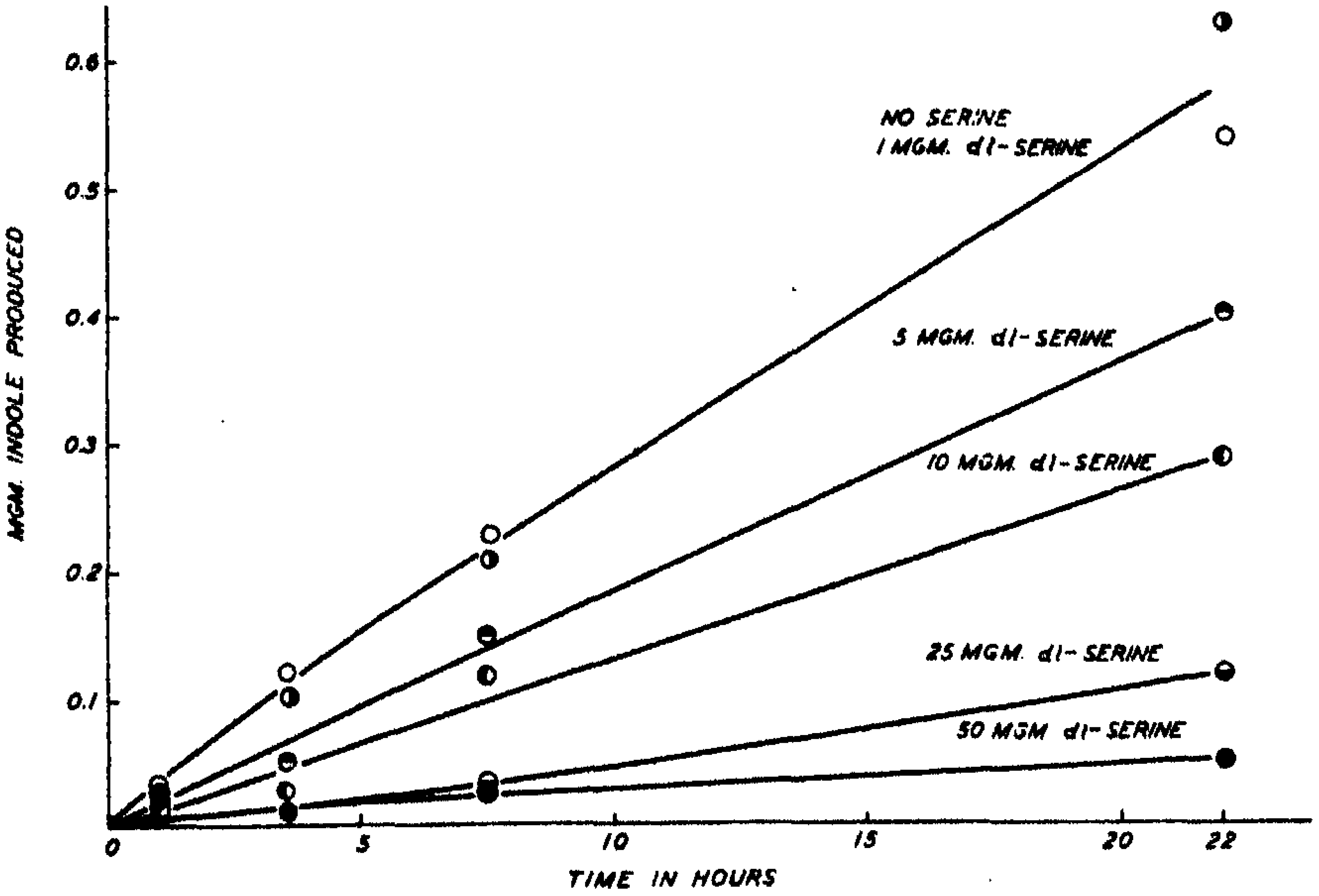


FIGURE 2

The effect of *dl*-serine on the production of indole from tryptophane by *E. coli* suspensions. Each 125-ml. culture flask contained 2 mg. *l*-tryptophane and 4 ml. of a suspension of washed *E. coli* cells (*ca.* 20 mg.) in 50 ml. of 0.05 *M* phosphate buffer at pH 7.2. The flasks were plugged with cotton and incubated at 25°C. with constant shaking.

These results support the view that the production of indole from tryptophane by *E. coli* actually involves a reversal of the reaction postulated for *Neurospora*. In this case indole and serine would be formed directly from tryptophane and the serine further oxidized, probably through pyruvic acid.[18] This oxidation of serine would account for the end-products of the reaction and the O_2 consumption which were observed by Woods.[8]

Krebs, Hafez and Eggleston[16] have suggested that *E. coli* produces indole from tryptophane by way of kynurenine and *o*-aminophenylacetaldehyde. In a direct test our active coli suspension failed to produce indole from l-kynurenine. The production of indole from *o*-aminophenylethanol in the experiments of Krebs, *et al.*, indicates only that the bacteria can oxidize the alcohol to the aldehyde. These postulated intermediates are probably not concerned in the normal production of indole from tryptophane, since our evidence supports the simpler hypothesis involving the direct production of indole and serine.

Discussion.—The β-hydroxyl group in serine apparently makes this amino acid quite reactive biologically. Binkley and du Vigneaud[17] have described the rôle of serine in the biosynthesis of cysteine from homocysteine. Chargaff and Sprinson[18] have studied the oxidation of serine by a number of bacteria including *E. coli*, and have suggested that the first reaction is an intramolecular dehydration with the formation of α-aminoacrylic acid. This reaction is analogous to the proposed intermolecular dehydration involved in the condensation of indole with serine.

The relation between indole and serine is not necessarily as simple as that proposed. The main direction of the reaction seems to be different in the two organisms used. Indole production is predominant with coli, but the reaction apparently is reversible. With *Neurospora* the synthetic reaction goes well, but it has so far been impossible to demonstrate the production of indole from tryptophane. However, analogous directional specificities are known in other reversible biological reactions such as the oxidation and reduction of inorganic nitrogen or sulfur compounds by bacteria.

The biosynthesis of tryptophane in *Neurospora* and probably in *E. coli* takes place by reactions which do not include the conventional reductive amination of the keto acid analog. This may be true of the synthesis of tryptophane in other organisms and possibly of the biosyntheses of certain other amino acids.

Summary.—The biosynthesis of tryptophane by the mold *Neurospora crassa* takes place through a direct reaction between indole and serine, possibly an intermolecular dehydration.

A reversal of this reaction is involved in the production of indole from tryptophane by the bacterium *Escherichia coli*.

---

* Work supported by grants from the Rockefeller Foundation.

[1] Fildes, P., *Brit. Jour. Exptl. Path.*, **22**, 293 (1941).

[2] Snell, E. E., *Arch. Biochem.*, **2**, 389 (1943).

[3] Beadle, G. W., and Tatum, E. L., *Proc. Nat. Acad. Sci.*, **27**, 499 (1941).

[4] Tatum, E. L., Bonner, D., and Beadle, G. W., *Arch. Biochem.* (in press).

[5] Tatum, E. L., and Bonner, D., *Jour. Biol. Chem.*, **151**, 349 (1943).

[6] Happold, F. C., and Hoyle, L., *Biochem. Jour.*, **28**, 1171 (1934).

[7] Generously made available by Dr. Max Bergmann, Rockefeller Institute, New York.

[8] Woods, D. D., *Biochem. Jour.*, **29**, 640 (1935).

[9] Determined by the pea test, see van Overbeek, J., and Went, F. W., *Bot. Gaz.*, **99**, 22 (1937).

[10] Berg, C. P., Rose, W. C., and Marvel, C. S., *Jour. Biol. Chem.*, **85**, 207 (1929–1930).

[11] du Vigneaud, V., and Sealock, R. R., *Ibid.*, **96**, 511 (1932).

[12] Microanalysis by Dr. A. J. Haagen-Smit, California Institute of Technology.

[13] Chargaff, E., and Sprinson, D. B., *Jour. Biol. Chem.*, **151**, 273 (1943).

[14] Evans, W. C., Handley, W. C. R., and Happold, F. C., *Biochem. Jour.*, **35**, 207 (1941).

[15] Krebs, H. A., Hafez, M. M., and Eggleston, L. V., *Ibid.*, **36**, 306 (1942).

[16] Culture obtained from Dr. C. E. Clifton, Department of Bacteriology, Stanford University.

[17] Binkley, F., and du Vigneaud, V., *Jour. Biol. Chem.*, **144**, 507 (1942).

A SECOND APPROXIMATION TO SOPER'S EPIDEMIC CURVE

BY EDWIN B. WILSON AND JANE WORCESTER

HARVARD SCHOOL OF PUBLIC HEALTH

Communicated January 20, 1944

If τ be the latent period between infection and infectiousness and σ be the period of infectiousness, the equation which expresses Soper's formulation of the course of an epidemic is[1]

$$A - \frac{dS}{dt} = rS[A\sigma - S(t - \tau) + S(t - \tau - \sigma)], \qquad (1)$$

where A is the rate at which new susceptibles are recruited to the population, S is the number of susceptibles at time t, and r is a rate of contact between the infectious and the susceptibles. The "law of mass action" is assumed, namely, that the rate at which the susceptibles are infected is proportional jointly to the number of susceptibles and the number of the infectious. If the right-hand member of (1) is expanded by Taylor's Theorem about the value $t = -(\tau + \sigma/2)$ we have

$$A - \frac{dS}{dt} = r\sigma S\left[A - \frac{dS}{dt}\bigg|_{t - \tau - \sigma/2} - \frac{\sigma^2}{24}\frac{d^3S}{dt^3}\bigg|_{t - \tau - \sigma/2} - \ldots\right]. \qquad (2)$$

Soper introduces a number of additional assumptions, of which one is that the time of infectiousness σ is so short that all the derivatives of higher order than the first may be disregarded.[2] For the steady state, could one exist, where S is constant and equal to m, equation (1) shows that $r\sigma S = 1$, that is, $r\sigma m = 1$. In that state one infectious person would infect just one susceptible. We may express S relative to this number m by introducing $x = S/m$. The equation then becomes

$$\frac{A}{m} - \frac{dx}{dt} = x\left[\frac{A}{m} - \left.\frac{dx}{dt}\right|_{t\,-\,\tau\,-\,\sigma/2}\right]. \tag{3}$$

The period $\tau + \sigma/2$ is Soper's "incubation period." If it be introduced as the unit of time, t must be replaced by $T(\tau + \sigma/2)$, and if the rate of recruits A be henceforth defined as the number of recruits in this unit of time, equation (3) will maintain the same form in T, viz.,

$$\frac{A}{m} - \frac{dx}{dT} = x\left[\frac{A}{m} - \left.\frac{dx}{dT}\right|_{T\,-\,1}\right]. \tag{3'}$$

We shall introduce the rate of new infections z measured relative to m and its logarithm, namely,

$$z = \frac{A}{m} - \frac{dx}{dT}, \qquad u = \log z,$$

in place of x. Then on differentiating (3′) and shifting the origin of time by a unit we have

$$\left.\frac{du}{dT}\right|_{T\,+\,1} - \left.\frac{du}{dT}\right|_{T} = -\left[1 - \frac{A}{m}e^{-u(T\,+\,1)}\right]e^{u(T)}, \tag{4}$$

which is the equation in u which corresponds rigorously to Soper's assumptions. By methods of approximation he arrives at the equation

$$\frac{d^2u}{dT^2} = -\left[e^u - \frac{A}{m}\right]. \tag{4'}$$

The rigorous derivation given above shows that he has effectively replaced on the left-hand side of the equation

$$\left.\frac{du}{dT}\right|_{T\,+\,1} - \left.\frac{du}{dT}\right|_{T} = \frac{d^2u}{dT^2} + \frac{1}{2}\frac{d^3u}{dT^3} + \frac{1}{6}\frac{d^4u}{dT^4} + \dots$$

by the first term of its expansion and on the right-hand side has replaced $e^{u(T)\,-\,u(T\,+\,1)}$ by 1 as though in this expression $u(T)$ and $u(T + 1)$ did not differ appreciably.

It is our purpose to examine, for the simple case where the epidemic is sufficiently short so that the number of recruits does not significantly affect its course, the integrals of

$$\frac{d^2u}{dT^2} = -e^u \qquad \text{and of } \left.\frac{du}{dT}\right|_{T\,+\,1} - \left.\frac{du}{dT}\right|_{T} = -e^{u(T)}. \tag{4''}$$

The integral of the first is readily found as

$$u = 2 \log \operatorname{sech} \sqrt{z_0/2}T + \log z_0, \qquad z = z_0 \operatorname{sech}^2 \sqrt{z_0/2}T, \quad (5)$$

where z_0 is the rate of new cases[3] at the peak of the epidemic which is taken as the origin of time. With $A = 0$, $z = -dx/dT$ and, if we integrate z, we have

$$x = \text{const.} - \sqrt{2z_0} \tanh \sqrt{z_0/2}T. \quad (6)$$

The constant may be determined if we know the value of x when $T = 0$, i.e., at the peak of the epidemic. The fundamental relations between C or z and S or x lead, however, in case $A = 0$, to the equation[4]

$$x(T) = \frac{z(T)}{z(T-1)} = \frac{\operatorname{sech}^2 \sqrt{z_0/2}T}{\operatorname{sech}^2 \sqrt{z_0/2}\,(T-1)}. \quad (6')$$

It is readily seen that (6) and (6′) are not equivalent no matter what value be assumed for the constant of integration. The discrepancy between (6) and (6′) cannot be avoided as it is inherent in the approximate nature of the solution in (6) for (4′). If we use (6′) we find that when $T = 0$, the value of x is $\cosh^2 \sqrt{z_0/2}$ and this value might be used for the constant in (6). On the other hand if we observe from (6′) that for $T = {}^1/_2$, $x = 1$, we should obtain the value of the constant in (6) as $1 + \sqrt{z_0/2} \tanh \sqrt{z_0/8}$ which is not equal to $\cosh^2 \sqrt{z_0/2}$ except as both may be considered equal to $1 + z_0/2$. Indeed if we substitute the approximation (5) in (4″), it is by no means clear that the approximation is good. Soper frankly made the hypothesis that the epidemics were so moderate that terms of higher order could be neglected[5]; it would follow that the discrepancies which we have been citing would have to be considered as inconsequential. Our aim is precisely to see whether a better approximation may not be found, at least in the simple case where the rate of recruits A may be neglected.

To discuss approximations one has to have some estimate of the values of z_0—how large are they? We cannot obtain z_0 by dividing reported cases at the peak during an incubation period by the value m, because in the first place the cases reported are only a fraction of the true cases and in the second place we have at best only crude estimates of the number m of susceptibles just sufficient for one case to make one new case. If we take a few clear-cut epidemics of measles we may determine the constants of the expression $z = A \operatorname{sech}^2 BT$, which is the first approximation to Soper's epidemic curve, either by a cumulative plot on growth paper or by a calculation based upon moments[6] using

$$B = \frac{\pi}{2\sqrt{3}\sigma}, \qquad A = \frac{\pi}{4\sqrt{3}\sigma} \times \text{total cases}, \quad (7)$$

where σ is the standard deviation of the time measured in fortnights (which

Soper takes to be the incubation period of measles). On the graph on growth paper[7] the value of the constant B is 2.945 divided by the elapsed time in incubation periods from the 5% to the 95% points on the fitted straight line. The value of B is independent of the fraction of cases reported, provided only that the fraction does not change during the course of the epidemic, whether it be determined graphically or by moments; and as, in the approximate solution, $B = \sqrt{z_0/2}$, the observed value of B gives a determination of z_0 in so far as the theory is valid. For a number of epidemics of measles we find (Fig. 1) values of B ranging[8] from about 0.20 to about 0.60.

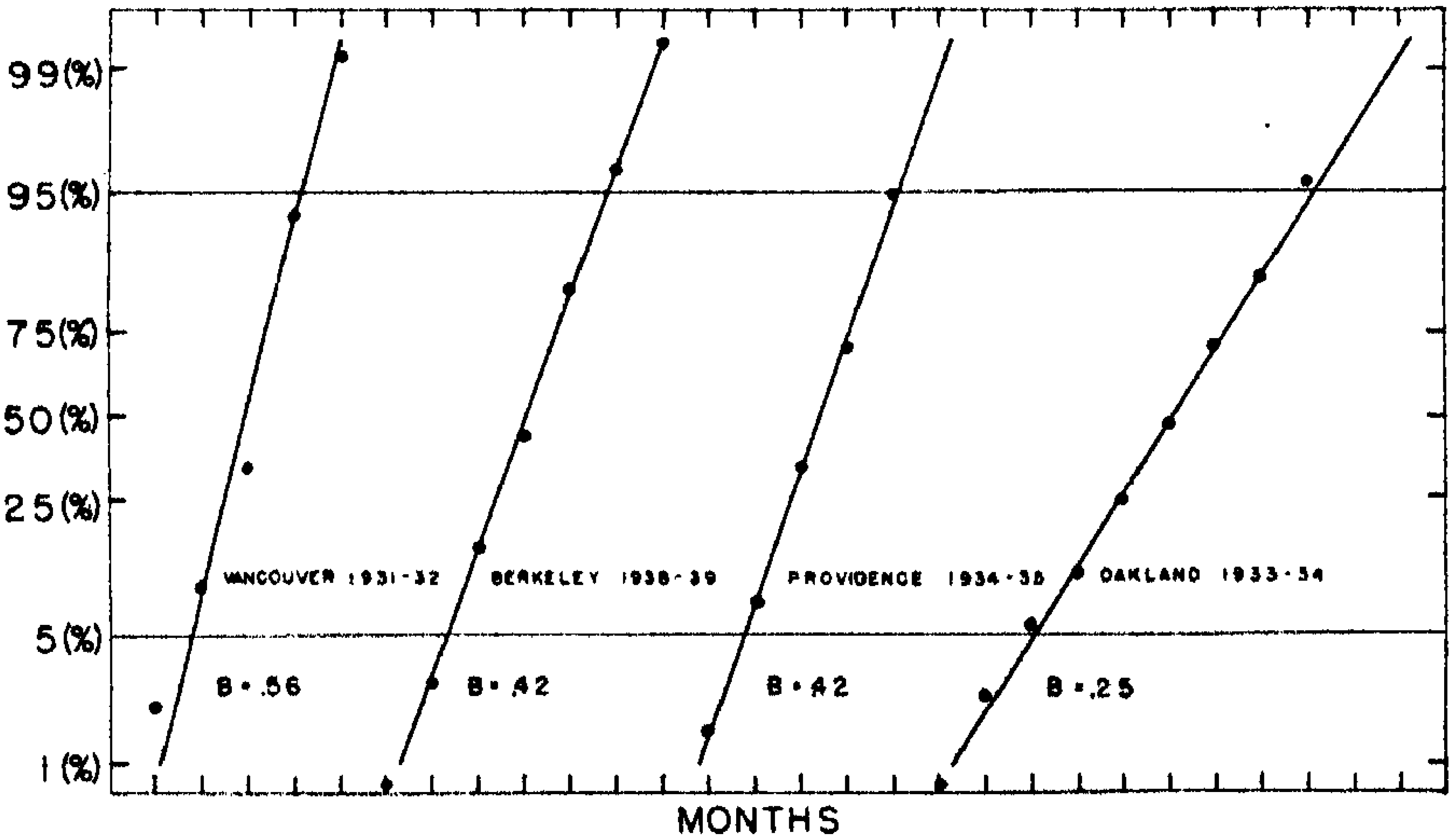


FIGURE 1

Cumulative graphs on growth paper of epidemics of measles in Vancouver, B. C., 1931–1932; Berkeley, Calif., 1938–1939; Providence, R. I., 1934–1935; and Oakland, Calif., 1933–1934.

The exact equation for the epidemic curve with $A = 0$, viz.

$$\left.\frac{du}{dT}\right|_{T+1} - \left.\frac{du}{dT}\right|_{T} = -e^{u(T)}, \tag{8}$$

gives the difference of the slopes a unit time apart. If it be considered that the values u_0 and $a = du/dT$ at any time T be known, the value of u at $T + 1$ upon the tangent to the curve will be $u_0 + a$; the tangent to the curve at $T + 1$ is, however, $a - e^{u_0}$ and if we had proceeded from T to $T + 1$ upon a line of this slope we should have reached a value $u_0 + a - e^{u_0}$. If we assume that a better value than either of these is their mean,[9] we have for the new value

$$u(T+1) = u(T) + \left.\frac{du}{dT}\right|_T - \frac{1}{2}e^{u(T)}. \tag{9}$$

With this relationship one can compute successive values of u stepwise from any assumed initial conditions $u = u_0$ and $du/dT = a$. To go backward stepwise it is necessary to solve for $u(T-1)$ the equation[10]

$$u(T-1) = u(T) - \left.\frac{du}{dT}\right|_T - \frac{1}{2}e^{u(T-1)}. \tag{9'}$$

In this manner we may plot out u as a function of the time for positive and negative integral values of T starting from any initial conditions, and may fill in intermediate values if we desire by using the parabola,

$$u(T+\theta) = u(T) + \left.\frac{du}{dT}\right|_T \theta - \frac{1}{2}e^{u(T)}\theta^2,$$

where θ is a positive proper fraction. A second method of following the epidemic curve would be to proceed by undetermined coefficients, writing

$$u = u_0 + a(T-T_0) + \beta(T-T_0)^2 + \gamma(T-T_0)^3 + \ldots \tag{10}$$

and determining β, γ, ... in terms of the initial conditions[11] u_0, a by substitution in (8). Terms up to and including that in $(T-T_0)^6$ seem to give an adequate solution to advance or retreat one time unit at a time when near the maximum and two or more units when fairly well removed from the maximum.

A different method of attack is to proceed directly to modify the equation $d^2u/dT^2 = -e^u$, which is the first approximation, over into one which is a better representation of (8). We have

$$\left.\frac{du}{dT}\right|_{T+1} - \left.\frac{du}{dT}\right|_T = \frac{d^2u}{dT^2} + \frac{1}{2}\frac{d^3u}{dT^3} + \frac{1}{6}\frac{d^4u}{dT^4} + \ldots = -e^u. \tag{11}$$

If we use the approximation $d^2u/dT^2 = -e^u$ to find d^3u/dT^3, we get as an equation of the second order

$$\frac{d^2u}{dT^2} - \frac{1}{2}e^u\frac{du}{dT} = -e^u.$$

This equation may be integrated in parametric fashion[12]; but it turns out that a much more convenient, and presumably better, approximation may be had by carrying the process one stage further on the original equation. If

$$\frac{d^2u}{dT^2} + \frac{1}{2}\frac{d^3u}{dT^3} = -e^u, \qquad \frac{1}{6}\frac{d^4u}{dT^4} = -\frac{1}{3}e^u\frac{du}{dT} - \frac{1}{3}\frac{d^3u}{dT^3},$$

and we may write, on substitution in (11),

$$\frac{d^2u}{dT^2} + \frac{1}{6}\frac{d^3u}{dT^3} - \frac{1}{3}e^u\frac{du}{dT} = -e^u.$$

The value of d^3u/dT^3, which would be obtained from the equation by differentiating the equation after neglecting the third derivative, may be obtained and substituted in the equation. The result is[13]

$$\frac{d^2u}{dT^2}\left(1 + \frac{e^u}{18}\right) - \frac{e^u}{2}\frac{du}{dT} + \frac{e^u}{18}\left(\frac{du}{dT}\right)^2 = -e^u. \qquad (12)$$

If $p = du/dT$, the first integral with $p = 0$ when $u = u_0$ is

$$p^2 + 12\left[\frac{18 + e^{u_0}}{18 + e^u} - 1\right]p - 36\left[\frac{18 + e^{u_0}}{18 + e^u} - 1\right] = 0.$$

This equation may be solved for p and integrated in the form

$$T = \frac{1}{6}\log\,(z/z_0) \pm \frac{1}{3}\sqrt{1 + 18/z_0}\,\tanh^{-1}\sqrt{1 - z/z_0}, \qquad (13)$$

where z_0 is the case rate at the peak of the epidemic and T is measured from that time, the + sign is used for $T > 0$ and the − sign for $T < 0$. Equation (12) may be solved by an iterative process[14] as

$$\log\frac{z}{z_0} = 2\,\mathrm{ls}\left\{aT - \frac{a}{3}\,\mathrm{ls}\left[aT - \frac{a}{3}\,\mathrm{ls}\,(aT - \text{etc.})\right]\right\}, \qquad (14)$$

where ls stands for log sech and a for $3/\sqrt{1 + 18/z_0}$.

TABLE 1

TABLE OF VALUES OF THE CASE RATE FOR SOME VALUES OF T

T	BY (5)	BY (9)	BY (10)	BY (14)
0	0.300	0.300	0.300	0.300
1	0.259	0.258	0.258	0.258
−1	0.259	0.263	0.262	0.261
2	0.174	0.168	0.167	0.166
−2	0.174	0.184	0.183	0.182
3	0.098	0.088	0.087	0.087
−3	0.098	0.112	0.110	0.110
4	0.050	0.041	0.040	0.040
−4	0.050	0.062	0.061	0.061
8	0.002	0.001	Not computed	0.001
−8	0.002	0.004	Not computed	0.004

The values of z, when the initial conditions are $z_0 = 0.3$ at maximum when $T = 0$, for the four approximations (5), (9), (10) and (14) are tabulated for some values of T (table 1). It is seen that the approximations (9), (10), (14) are very much alike and that they give an asymmetrical curve which differs considerably from the symmetrical curve (5). The skewness of the curve is negative, the rise to the peak being slower than the fall from it. For the particular data tabulated this skewness, as measured by the ratio of the third moment to the cube of the standard deviation, is about -0.30. The formula (13) or (14) represents the epidemic curve that follows from Soper's theory to a much higher degree of approximation than (5) and, we believe, to a quite sufficient accuracy for values of z_0 up to 0.4, that is, for epidemics so severe that at their peak the number of new cases per incubation period is about four-tenths of the number of susceptibles just sufficient for each old case to generate one new case.

---

[1] Wilson, E. B., and Burke, Mary H., these PROCEEDINGS, **28**, 361–367 (1942), footnote 2. The reference to Soper, H. E., is *Jour. Roy. Statist. Soc. (London)*, **92**, 34–61 (1929). If the rate of new infections is $C(t)$ and the number of infectious persons is $I(t)$ the equation $-dS/dt = C(t) - A(t)$ implies that none of the persons who become infected return to the population of susceptibles unless allowance therefor is made in the rate of recruits A. The "law of mass action" is $C = rIS$ with $I = \int_{t-\tau-\sigma}^{t-\tau} C(t)dt$. The rate A is taken as constant.

[2] We keep $\sigma/2$ in the expression $\tau + \sigma/2$ because it is possible that σ be small enough so that the higher derivatives in (2) may be neglected and yet large enough to make $\tau + \sigma/2$ a somewhat more accurate "incubation period" or period between generations of the infected than τ.

[3] The rate of new cases $z(T) = C(T)/m$ is the rate at which susceptibles become infected, not the rate at which persons become clinical cases, which is presumably more nearly equivalent to $z(T + 1)$.

[4] For, when σ is small $I = \int_{t-\tau-\sigma}^{t-\tau} C(t)dt = \sigma C(t - \tau - \sigma/2) = \sigma C(T - 1)$ and $C = rIS$ becomes $C = r\sigma SC(T - 1)$ which on division by m with $r\sigma m = 1$ leads to $z(T) = xz(T - 1)$.

[5] It should be remarked that Soper's main interest was in periodicity, which on the present theory requires that the rate of recruits A should not be zero, and which brings in additional difficulties; we have no desire to minimize Soper's accomplishments in his paper.

[6] When using moments it is particularly desirable that the epidemics rise from 0 and drop to 0 in a clear-cut fashion because irregularities at the ends may have a considerable effect on the moments; effects due to irregular beginning and ending can be more readily disregarded when using the graphical method.

[7] Wilson, E. B., these PROCEEDINGS, **11**, 451–456 (1925).

[8] If we use the curve $A \operatorname{sech}^2 BT$ to describe the epidemic empirically (or any similar bell-shaped curve), the values of A and B may be regarded as determinable independently from the data, but as the integral of the approximate differential equation is $z = z_0 \operatorname{sech}^2\sqrt{z_0/2}\,T$, it follows that A and B are connected by the relation $A = 2B^2$ and that total cases relative to m are $2A/B$ or $4B$. The range of total cases would therefore be from about $0.80m$ to about $2.40m$ in these epidemics in so far as the theory is valid.

[9] This is equivalent to fitting a parabola. For if $u = a + bT + cT^2$,

$$\left.\frac{du}{dT}\right|_T = b + 2cT, \quad \left.\frac{du}{dT}\right|_{T+1} = b + 2cT + 2c, \quad \left.\frac{du}{dT}\right|_{T+1} - \left.\frac{du}{dT}\right|_T = 2c = -e^{u(T)}$$

give the result (9) for $u(T + 1) = a + bT + cT^2 + b + 2cT + c$.

[10] If we first neglect $\frac{1}{2}\, e^{u(T-1)}$ we find $u(T-1) = u(T) - (du/dT)_T$ which may be substituted in $\frac{1}{2}\, e^{u(T-1)}$, and so on. This iterative process converges rather rapidly.

[11] We find

$$\beta = -\frac{e^{u_0}(a^2 - 6a + 12)}{2(12 + e^{u_0})}, \qquad \gamma = -\frac{2}{3}\,\beta\,\frac{(ae^{u_0} - 3e^{u_0} - 6a)}{12 + e^{u_0}}$$

$$\delta = \frac{e^{u_0}}{24}\left(2\beta - 3a\beta - 3\gamma + a^2 - \frac{1}{2}\,a^3\right), \qquad \epsilon = -\frac{e^{u_0}}{20}\left(\gamma + a\beta + \frac{1}{6}\,a^3\right).$$

[12] Ans.:

$$u = \log\left[e^{u_0} + 4\log\left(1 - \frac{1}{2}\,p\right) + 2p\right],$$

$$T = \int_0^p \frac{-dp}{\left(1 - \frac{1}{2}\,p\right)\left[e^{u_0} + 4\log\left(1 - \frac{1}{2}\,p\right) + 2p\right]}.$$

[13] Various other equations may be obtained by similar methods of approximation but this one seems to be the one which gives the least inconvenient integrals.

[14] The convergence of this iterative formula is fairly rapid; if one knows in advance an approximate value for $\frac{1}{2}\log(s/s_0)$ at time T and takes this value as ls (aT − etc.), the convergence will be more rapid than if one starts with "− etc." as zero and computes ls aT as the first step.

ERRATA

In *Proc. Nat. Acad. Sci.*, **28**, 374–377 (1942), the following corrections are needed:

In (1.1) $sh(x)$ should be $sh(\pi x)$.

In (1.3) the factor $^1/_2$ should be $(1/2i)$.

In (3.3) a factor $m\pi$ cosec $(m\pi)$ should be inserted on the right.

On account of the last correction the orthogonal relation for $E_n(x)$ can be found. It is

$$\int_{-\infty}^{\infty} E_n(ix)E_{n'}(-ix)dx/sh^2(^1/_2\pi x) = 0 \qquad n' \neq n,\ n' > 0,\ n > 0$$
$$= 2/[\pi n(n + {}^1/_2)(n + 1)] \qquad n' = n > 0$$

In *Proc. Nat. Acad. Sci.*, **28**, 371–374 (1942):

On p. 371 the equation

$$\lim_{n \to \infty} S_n^{(r)} = 1$$

should be accompanied with the restriction $|1 - x^{-1}| > 1$. If $|1 - x^{-1}| = 1$ the condition $R(m + r) > 1$ should be imposed.

PROCEEDINGS
OF THE
NATIONAL ACADEMY OF SCIENCES

Volume 30 March 15, 1944 Number 3

Copyright 1944 by the National Academy of Sciences

RESPIRATORY RATES OF THE SHOOT TIPS AND MATURING TISSUES IN LUPINUS ALBUS AND TROPAEOLUM MAJUS

BY ERNEST BALL* AND EDGAR J. BOELL

OSBORN BOTANICAL LABORATORY AND OSBORN ZOÖLOGICAL LABORATORY, YALE UNIVERSITY

Communicated February 3, 1944

It has long been recognized that the shoot apex in the higher plants serves as the region from which new tissues of the primary plant body are produced.[1-9] Indeed, Bower[2] has characterized the activities of this meristem as an indefinitely continued embryology.

The work of several authors[12-15] suggests that the shoot apex may influence, to some extent at least, the differentiative activity of subjacent tissues in the plant, for when shoot apices are regenerated in mature parenchymatous tissues, the new shoot apices apparently induce the formation of vascular strands which connect with the original conductive system of the shoot. In somewhat the same way the so-called organizers in the animal embryo have been demonstrated to influence the development and differentiation of other regions. Accordingly it would be of interest to determine whether the metabolic characteristics of the organizer[19] and of the plant shoot apex are to some extent similar.

Furthermore, in view of the widely current idea, originally enunciated by Pfeffer,[16] that growth is usually accompanied by elevated respiratory activity, it is especially important to study the relationship between respiration and growth in the shoot apex. The embryonic region exhibits the highest growth rate of the shoot on the basis of the following criteria of growth—frequency of cell division, synthesis of protoplasm, and changes in shape and structure associated with the production of foliar primordia—and if Pfeffer's thesis is correct, it therefore ought to possess the greatest respiratory activity.

Some experimental work has previously been done on the respiration of various gross regions along the developing plant axis. Thus Kidd, West and Briggs[10] found the liberation of carbon dioxide in the sunflower to be greater in the bud than in more basal tissues. Caldwell and Meiklejohn[11] measured the respiration of the entire terminal bud and of the subjacent tissues in the tomato and found a gradient of decreasing oxygen consumption from apex to base. However, in order to gain an idea of the significance of respiratory rate in the phenomena of induction, production of foliar primordia and maturation, as well as in the very rapid growth of the shoot apex, still more precise techniques must be applied to very small volumes of tissues. It is necessary to measure separately and with great accuracy the oxygen consumption of the shoot apex (i.e., of the embryonic zone) and of the subjacent regions in which cellular maturation is occurring

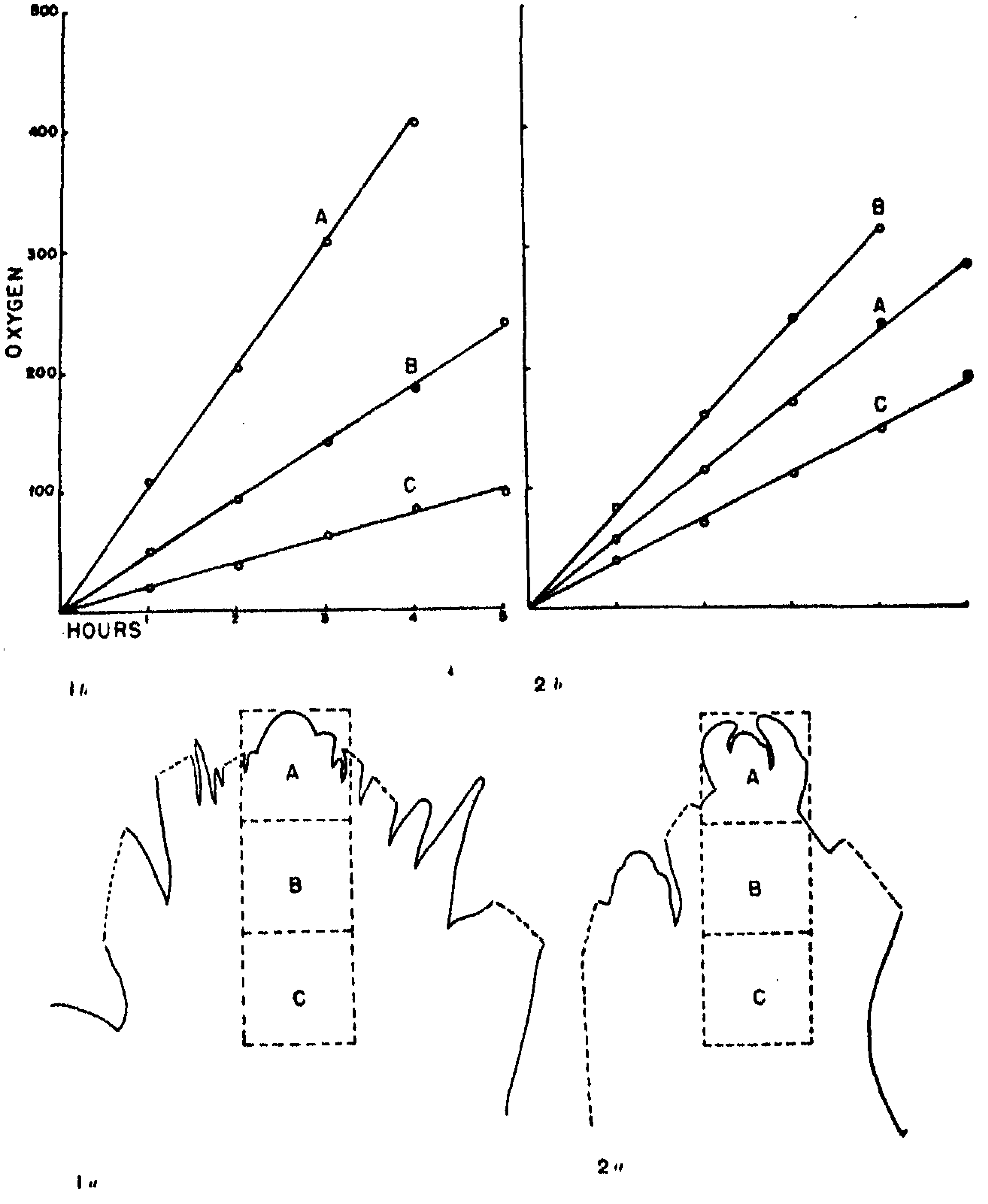
In the present work a comparative survey has been made of the rate of oxygen consumption of the meristematic and of the immediately subjacent maturing tissues in the terminal part of the plant axis, in order to determine the relationship between the respiratory activity of a given region and the degree of its cellular maturation, or its position in the gradient of differentiation.

Materials and Methods.—Young plants of *Lupinus albus* L. and of *Tropaeolum majus* L. were used. *Lupinus* shows strong apical dominance, and branching by the growth of axillary buds is rare. *Tropaeolum*, on the other hand, exhibits apical dominance only during the early seedling stages, and the growth of shoots from all axillary buds soon leads to a much branched condition. Thus, the present work was done upon plants of two different physiological types.

Pieces of tissues, designated as *A*, *B* and *C* in the order of their position down the shoot (Figs. 1 (*a*) and 2 (*a*)) and each having a volume of about 430 μ^3 (ca. 0.08 mm.^{3}), were excised from the shoot tip under the high power of a binocular dissecting microscope equipped with a calibrated ocular. The *A* piece comprised the shoot apex and the three youngest visible foliar primordia. In addition, a small amount of non-apical tissue was also unavoidably included, but since the volume of the latter was insignificant in comparison with that of the apical tissue proper, the oxygen consumption of the *A* piece, for all practical purposes, can legitimately be considered typical of the actual apical tissue. Piece *B* was composed of maturing tissue immediately basal to the *A* piece, while piece *C* represented still further maturing tissue subjacent to region *B*.

The oxygen consumption of the tissues was determined by means of the Cartesian diver technique.[17, 18] Each of the pieces of tissue was placed in the bulb of a diver with 1.5 μl. of Pfeffer's culture solution (Robbins[20]), and readings of oxygen consumption were begun, and continued for periods of four to five hours. During the experiments, the apparatus was darkened in order to prevent photosynthesis in the chlorenchymatous tissues of *Lupinus*.

Results and Discussion.—The results of typical experiments on *Lupinus* and *Tropaeolum* are shown in figures 1 (*b*) and 2 (*b*). The graphs indicate



FIGURES 1 AND 2

Figures 1 (*a*) and 2 (*a*) are line drawings of median sections of stem tips of *Lupinus albus* and of *Tropaeolum majus*, respectively, and show the amounts and locations of the *A*, *B* and *C* pieces of tissues.

Figures 1 (*b*) and 2 (*b*) show the oxygen consumption of the *A*, *B* and *C* pieces of tissues in the Cartesian divers. Ordinates denote mμl. of O_2 consumed per piece of tissue (ca. 0.08 mm.^{3}), and abscissas indicate length of time pieces of tissue were in the divers.

Figure 1 (*b*) indicates that there is a decreasing respiratory gradient along the shoot of *Lupinus*. Figure 2 (*b*) indicates that the subjacent (*B*) tissues in the shoot of *Tropaeolum* consumed more oxygen than the shoot tip.

that the rate of oxygen consumption of the excised tissues is essentially uniform during the period of measurement. It is significant that in none of the experiments was there any appreciable acceleration or depression of respiration during the initial phase of the experimental period, because this fact shows that the tissues were neither perceptibly stimulated nor depressed by the processes of excision from the plant body and immersion in the culture fluid. The general linearity of the curves, moreover, attests to the fact that nutritive conditions in the medium were such as to maintain the tissues in an essentially normal physiological state.

The curves reveal significant differences in the rate of oxygen consumption in the various regions of the same plant. Moreover, they show that the respiratory rate of a given region in relation to other regions is different in the two species studied. Thus, in *Lupinus*, a gradient in the rate of respiration decreasing from tip to base seems clearly indicated; while in *Tropaeolum* this is not the case, the most apical portion of the shoot tip (*A*) having a lower rate than the region of maturing cells immediately subjacent. However, in both plants the most basal portion of the shoot tip (*C*) shows the lowest rate of respiration.

Histological examination of the shoot tips of *Lupinus* and of *Tropaeolum* reveals marked differences in the amount of protoplasm in the regions selected for study. Many cells (especially in the shoot apex) of the *A* piece are densely cytoplasmic; those in region *B* are characterized by the beginning of conspicuous vacuolization and the appearance of intercellular spaces, and those in region *C* show still more prominently these indications of cellular maturation. It is highly probable, therefore, that the amount of actively metabolizing material in the three regions is not identical in spite of their volume equivalence. But that the differences in respiration reported above are not due altogether to differences in the amount of actively respiring material per unit of volume is shown by the fact that when respiratory rate is calculated on a dry weight basis the ratio of the respiratory activity in the three regions remains essentially the same (table 1). Fundamental differences in the respiratory rates of the three regions thus seem clearly indicated.

The quantitative differences in the dry weights of the *A*, *B* and *C* pieces, shown in table 1, are reflected also in preliminary data on the total nitrogen content of these regions. Although, as the table reveals, the distribution of total nitrogen does not altogether parallel the dry weight, the respiratory rates when calculated on the basis of total nitrogen show, in general, the same fundamental differences among the regions, and between the two plants.

In *Lupinus*, the respiratory rate of the shoot tip is higher than that of the subjacent tissues while in *Tropaeolum* this appears not to be true. In both *Lupinus* and *Tropaeolum* the shoot apex serves as the region from which new foliar primordia and new stem tissues are produced, and in both plants, moreover, the apical region displays the highest growth rate of the shoot. It is evident therefore, in view of the data presented, that these activities of the shoot apex are not necessarily correlated with a high respiratory rate. A comparable situation is seen in amphibian development where it appears that the organizer activities of the dorsal lip of the blastopore are not correlated with respiratory level.[19] It may be of some significance that the respiratory activity of the various zones of *Lupinus*, in which apical dominance is manifest, is different from that in *Tropaeolum* where apical dominance is absent.

TABLE 1

REGION	NO. OF EXPERIMENTS	MμL. O_2 CONSUMED PER HOUR PER PIECE OF TISSUE (AVERAGE)	μG. DRY WEIGHT PER PIECE (AVERAGE)	MμL. O_2 CONSUMED PER HOUR PER μG. DRY WT. (AVERAGE)	μG. N PER PIECE (AVERAGE)	MμL. O_2 CONSUMED PER HOUR PER μG. OF NITROGEN (AVERAGE)
			Lupinus			
A	11	96	11.0	8.7	0.8	120
B	11	41	9.1	4.5	0.32	128
C	11	22	8.8	2.5	0.45	49
			Tropaeolum			
A	7	57	12.3	4.6	0.89	64
B	7	79	10.5	7.5	0.91	87
C	7	38	8.2	4.7	0.70	55

Summary.—Using the Cartesian diver technique, a study has been made of the respiratory activity in the shoot tips and subjacent tissues of *Lupinus albus* and *Tropaeolum majus*. The results indicate that differences in the distribution of respiratory activity along the apical-basal axis of the shoot exist in the two plants, but that there is apparently no correlation between the developmental activities of a given region and the magnitude of its respiratory rate.

---

[*] National Research Council Fellow in Botany, 1941–1942. Present address: Biological Laboratories, Harvard University.

[1] Wolff, C. F., *Theoria generationis*, Halae, 1759.

[2] Bower, F. O., *Botany of the Living Plant*, London, 1919.

[3] Child, C. M., *Jour. Exp. Zool.*, **11**, 187–220 (1911).

[4] Child, C. M., *Individuality in Organisms*, Chicago, 1915.

[5] Child, C. M., *Biol. Bull.*, **39**, 147–187 (1920).

[6] Child, C. M., *Protoplasma*, **5**, 447–476 (1928).

[7] Child, C. M., *Patterns and Problems of Development*, Chicago, 1941.

[8] Foster, A. S., *Bull. Torrey Bot. Club*, **68**, 339–350 (1941).

[9] Thoday, D., *Nature*, **144** (3648), 571–575 (1939).

[10] Kidd, F., West, C., and Briggs, G. E., *Proc. Roy. Soc. Lond.* (*B*), **92**, 368–384 (1921).

[11] Caldwell, J., and Meiklejohn, J., *Ann. Bot.*, N. S., **1**, 477–486 (1937).

[12] Simon, S., *Jahrb. wiss. Bot.*, **45**, 351–478 (1908).

[13] Hammond, B. L., *Bot. Gaz.*, **97**, 834–845 (1936).

[14] Naylor, E. E., and Johnson, B., *Amer. Jour. Bot.*, **24**, 637–678 (1937).

[15] McVeigh, I., *Ibid.*, **25**, 7–11 (1938).

[16] Pfeffer, W., *The Physiology of Plants*, Trans. by A. J. Ewart, Oxford, 1903.

[17] Boell, E. J., and Needham, J., *Proc. Roy. Soc. Lond.* (*B*), **127**, 363–378 (1939).

[18] Boell, E. J., and Woodruff, L. L., *Jour. Exp. Zool.*, **87**, 385–402 (1941).

[19] Boell, E. J., *Growth*, Supplement to Vol. 6, pp. 37–53 (1942).

[20] Robbins, W. J., *Amer. Jour. Bot.*, **29**, 241–245 (1942).

---

*INVALIDATION OF THE CYTOLOGICAL EVIDENCE FOR RECIPROCAL CHIASMATA IN THE SEX CHROMOSOME BIVALENT OF MALE DROSOPHILA**

By Kenneth W. Cooper

Department of Biology, Princeton University

Communicated February 7, 1944

Darlington's (1934)[1] interpretation of meiosis in male *Drosophila pseudoobscura* involves a mechanical dualism in the modes of conjunction of autosomes and sex chromosomes. Thus autosomes are held to conjoin as bivalents by exaggeration of the forces normally expressed in somatic pairing,[2] whereas sex chromosomes are believed to conjoin solely by chiasmata. The primary evidence which led Darlington to this interpretation is purely cytological, and the more critical points may be resuméed as follows.

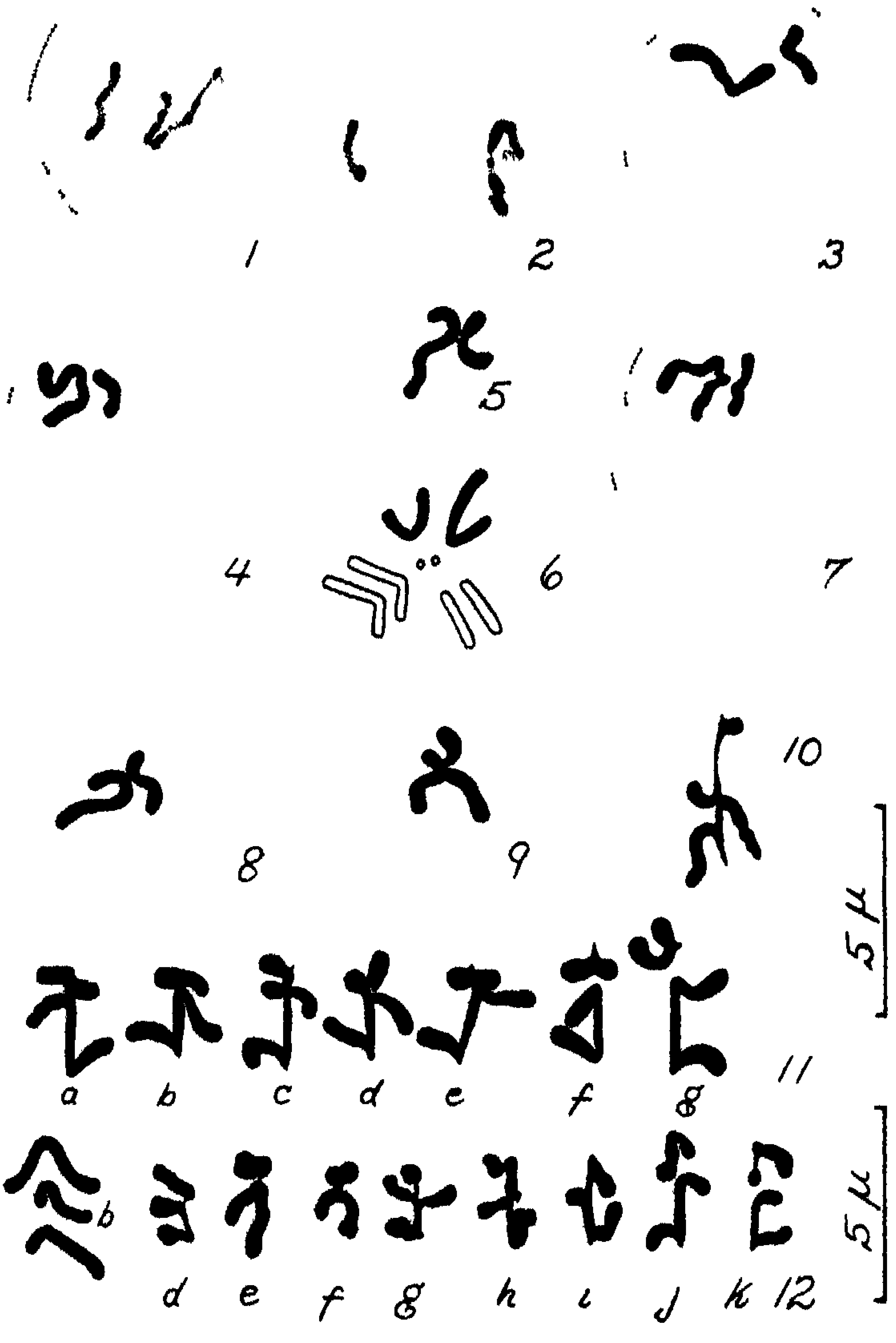
(1) Autosomal bivalents have homologues paired throughout their lengths at late diakinesis, *neither contacts* nor chiasmata being visible between them.

(2) In contrast to the autosomal bivalents, from diakinesis to onset of anaphase the sex chromosomes are *in contact* over a short segment of their lengths, and manifestly this is not the whole of their homologous regions. Furthermore, lengths of sex chromosomes which are not conjoined lie part in spite of their homology (Figs. 12 (*d*–*f*), diakinesis; Figs. 12 (*g*–*i*), metaphase-I).

(3) At prometaphase and metaphase there appears to be a state of tension between the kinetochores and the points of association in the sex chromosomes (Figs. 12 (*g*–*i*)).

(4) *X* and *Y* are unchanged when they separate at anaphase (Figs. 12 (*j*–*k*)). Were a single exchange present between the interstitial segments of contact and conjunction, the morphology of the separating chromosomes would clearly disclose it because *X* and *Y* are heteromorphic.

The striking contrast of the sex chromosome bivalent with the chiasma-free autosomal bivalents led Darlington to suspect that fundamentally different mechanisms are in operation in the two instances. According to



FIGURES 1–12

Comparison of sex chromosome conjunction in the spermatogenesis of *Olfersia bisulcata* Macq and of *Drosophila pseudoobscura* Frolova Except for figure 6 only sex chromosomes are illustrated, figures semidiagrammatic

Figure 1, late diplotene. Figures 2–4, 7, early diakinesis Figure 5, late diakinesis or early prometaphase Figure 6, metaphase plate of male gonial cell (sex chromosomes black) Figure 8, late diakinesis. Figures 9, 10 prometaphase. Figures 11 (*a–g*), sex chromosome bivalent of *Olfersia* at metaphase-I (*a–e*) and early anaphase I (*f, g*) Figures 12 (*a–k*), sex chromosomes of *Drosophila pseudoobscura*, redrawn from Darlington (1934)[1] Figure 12 (*a*), the *X*-chromosome. Figure 12 (*b*), *Y*-chromosome of weak-A race Figure 12 (*c*), *Y*-chromosome of weak-B race Figure 12 (*d–f*), diakinesis, weak-A race Figures 12 (*g–i*) metaphase I, weak-A races. Figures 12 (*j, k*) anaphase-I, weak-B races

Upper scale, approximate magnification for figures of *Olfersia* Lower scale, approximate magnification for figures of *Drosophila*.

his chiasma hypothesis of metaphase pairing[3] chiasmata provide the primary if not only means of conjunction of the chromosomes as bivalents. Observations (2), (3) and (4) above are compatible with the known properties of bivalents having chiasmata. Darlington[1] thus regards them as affording a demonstration of the presence of chiasmata in the sex chromosome bivalent. Since *genetic* crossing-over between the sex chromosomes of the male of *Drosophila* is rare and probably not a meiotic occurrence, the hypothesized chiasmata must be considered localized in inert regions and invariably to occur in reciprocal pairs. This, in brief, is the nature of the evidence for the reciprocal chiasmata hypothesis of sex chromosome conjunction in male *Drosophila*.

Short, localized, terminal pairing-segments have been discovered in the fly *Melophagus ovinus*.[4] No chiasmata need be assumed to conjoin *Melophagus'* autosomes in these segments at meiosis in the male; yet the autosomal bivalent configurations and behavior closely parallel those of the sex chromosome bivalent in male *Drosophila*.[5] For this reason it was likewise concluded that there is no cytological need to suppose that reciprocal chiasmata conjoin the sex chromosomes of *Drosophila*. The truth of this conclusion is admirably demonstrated by the details of bivalent formation by the sex chromosomes in the male of the fly *Olfersia bisulcata* Macq., a member of the same family to which *Melophagus* belongs (Hippoboscidae).

The diploid chromosome set of male *Olfersia bisulcata* ($2n = 8$) is remarkably different from that of *Melophagus ovinus* ($2n = 18$). There is a pair of submedian autosomes, a pair of nearly terminal autosomes, a pair of dot-like autosomes, and a heteromorphic pair of submedian sex chromosomes (Fig. 6; compare the sex chromosomes with those of *Drosophila pseudoobscura*, Figs. 12 (*a*, *b*, *c*)).[6]

At late diplotene the sex chromosomes of *Olfersia* may easily be recognized; as in *Melophagus* they are neither paired nor even close to one another (Figs. 1, 2). Henceforth, however, their behavior is unlike that of the sex chromosomes of *Melophagus* which remain unpaired, for X and Y ultimately conjoin in *Olfersia*. It will be seen that the resulting sex chromosome bivalent is identical with that of male *Drosophila pseudoobscura*.

As diakinesis progresses the sex chromosomes move closer together. During this movement an interstitial segment near the kinetochore in each chromosome becomes juxtaposed and finally closely conjoined (Figs. 3–4, 7–8, 5). At prometaphase the sex chromosome bivalent thus formed orients on the spindle (Fig. 9). During this process the regions between the kinetochores and the points of association may become markedly attenuated and hence appear to be in a state of tension (Fig. 10).[7] At late diakinesis, prometaphase and metaphase, the sex chromosomes of male

Olfersia are in contact only over a short interstitial segment of their lengths, and the lengths of the sex chromosomes which are not thus conjoined lie apart (Figs. 5, 8, 9, 10, 11 (*a–e*) of *Olfersia;* compare with Figs. 12 (*d–i*) of the sex chromosome bivalent of *Drosophila pseudoobscura* at the same stages).[8] Furthermore, the sex chromosomes of *Olfersia* are unchanged when they separate at anaphase (Figs. 11 (*f–g*) of *Olfersia;* compare with Figs. 12 (*j–k*) of *Drosophila pseudoobscura*).[9] Thus the sex chromosome bivalent of male *Olfersia* shows the same morphological attributes as are said to characterize the X-Y bivalent of *Drosophila.* But inasmuch as the sex chromosomes of male *Olfersia* conjoin *after* condensation, the resulting bivalent must be devoid of chiasmata.[10] Therefore none of the aforementioned attributes, singly or in combination with each other, can be considered to afford a demonstration of the presence of chiasmata. Accordingly it is concluded that there is no longer a cytological need to suppose the existence of reciprocal (or any other) chiasmata in the inert regions of the sex chromosomes of *Drosophila.* Just as in the cases of the autosomes of *Melophagus* and *Olfersia,* and the sex chromosomes of *Olfersia,* the X-Y bivalent of male *Drosophila pseudoobscura* may be supposed to be conjoined by means of localized pairing segments which do not form chiasmata.

It is not known whether the sex chromosomes in *Drosophila pseudoobscura* conjoin after or during their completion of condensation. That in fact they may actually conjoin while condensing is suggested by Darlington's[1] observation that at diakinesis not only are some of the sex chromosome bivalents visibly conjoined, but another configuration of sex chromosomes is also found. In this latter case the sex chromosomes are said to show a clear space between them without any visible connection joining the two chromosomes. Until *Drosophila pseudoobscura* is restudied the exact significance of this fact cannot be judged, but it is not improbable that careful seriation will show that the connected bivalents are a consequence of pairing after or during condensation as has been found to be the case in *Olfersia.* Such delayed pairing following condensation is not unique. There are many instances known of similar chromosomal maneuvers in other insects, some of which Schrader[11] has briefly reviewed. The entire problem of bivalent formation and modes of conjunction requires reanalysis, for the chiasma hypothesis of metaphase pairing is clearly incapable of generalizing all of the facts.

* This work was supported by a grant from the Penrose Fund of the American Philosophical Society, to which organization grateful acknowledgment is made.

[1] Darlington, C. D., *Genetics,* **19,** 95–118 (1934).

[2] This view is implicit in Steven's (*Jour. Exp. Zool.,* **5,** 359–374 (1908)) classic and initial account of spermatogenesis in male flies, and underlies my interpretation of the *Melophagus* case as well. Bauer (*Fortschr. Zool.,* **4,** 584–597 (1939)) and Wolf (*Chromosoma,* **2,** 192–246 (1941)) have questioned the validity of such an interpretation of autosomal conjunction at meiosis. Their arguments will be examined elsewhere.

[3] Darlington, C. D., *Jour. Gen.*, **21**, 17–56 (1929).

[4] Cooper, K. W., *Proc. Nat. Acad. Sci.*, **27**, 109–114 (1941).

[5] The only striking difference is that conjunction is terminal in *Melophagus* and interstitial in the sex chromosomes of *Drosophila*.

[6] Although the meiotic autosomal bivalents of *Olfersia* can be shown to conjoin by conjunctive segments devoid of chiasmata as in *Melophagus*, they will not be discussed here nor be illustrated. Suffice it to say that the bivalent formed by the submedian pair of autosomes is more complicated in structure than earlier believed (Cooper, K. W., *Amer. Phil. Soc. Yearbook, 1941*, pp. 122–127, (1942)), and only in the disjunction figures does it resemble autosomal bivalents of *Drosophila*.

[7] Hence identical with point (3), page 50.

[8] Hence identical with point (2), page 50.

[9] Hence identical with point (4), page 50.

[10] The extensive cytological data which show crossing-over to be a zygotene or pachytene event make implausible the assumption that crossing-over may occur between markedly condensed, post-pachytene chromosomes.

[11] Schrader, F., in: *Cytology, Genetics and Evolution*, University of Pennsylvania Press, Philadelphia, 1941, pp. 27–36.

BIRKHOFF'S THEORY OF GRAVITATION AND EINSTEIN'S THEORY FOR WEAK FIELDS

By Alberto Barajas

Institute of Mathematics, National University of Mexico

Communicated February 9, 1944

Recently, in the widely read *Mathematical Reviews*,[1] Weyl has commented on Birkhoff's theory of gravitation of 1942,[2] suggesting that it is much the same as a limiting form of Einstein's well-known theory of 1916 for weak fields, characterized by Weyl as follows: "If one writes the equations of Einstein's theory of gravitation for g_{ik} differing but infinitely little from the constant values of δ_{ik} characteristics for flat space-time, $g_{ik} = \delta_{ik} + h_{ik}$, and neglects higher powers of h_{ik}, one obtains a theory with a symmetric tensor of gravitation h_{ik} in a flat world, which in its factual consequences agrees almost completely with Einstein's original theory." I propose to show[3] that this degenerate form is fundamentally different from Birkhoff's theory and is unsatisfactory from a physical point of view. In doing so, attention will be limited to gravitational fields in empty space.

Birkhoff's space-time is the same as Minkowski's with $ds^2 = dt^2 - dx^2 - dy^2 - dz^2$. His fundamental gravitational tensor h_{ij} satisfies the equation $\Box h_{ij} = 0$ in empty space, and the equations of motion of a free particle in an arbitrary gravitational field are:

$$\frac{d^2x^i}{ds^2} = \delta^{pi}\left(\frac{\partial h_{jp}}{\partial x^k} - \frac{\partial h_{jk}}{\partial x^p}\right)u^ju^k$$

where $\delta^{pi} = 1$ for $p = i = 1$, $\delta^{pi} = -1$ for $p = i = 2, 3, 4$, zero otherwise, and $u^i = dx^i/ds$.

In the general relativity theory of Einstein the quadratic fundamental form may be written $ds^2 = (\delta_{ij} + H_{ij})dx^idx^j$. For weak gravitational fields, it is readily shown that H_{ij} may be chosen so that Einstein's gravitational equation $R_{\mu\nu} = 0$ reduces to $\square H_{ij} = 0$.[4] The equations of motion of a free particle are, for this degenerate case:

$$\frac{d^2x^i}{ds^2} = -g^{pi}\left(\frac{\partial H_{jp}}{\partial x^k} - \frac{1}{2}\frac{\partial H_{jk}}{\partial x^p}\right)u_ju^k$$

In this expression, *ds* is, of course, the interval in a curved space-time, not Minkowski's *ds*, although Weyl (loc. cit.) considers the degenerate theory to have reference to "a flat world."

If we consider the static field due to a mass-point *m*, for reasons of symmetry, H_{ij} must have the form (r = radial distance):

$$H_{ij} = \begin{Vmatrix} a/r & 0 & 0 & 0 \\ 0 & b/r & 0 & 0 \\ 0 & 0 & b/r & 0 \\ 0 & 0 & 0 & b/r \end{Vmatrix}$$

with $a = -2m$ so as to obtain the classical Newtonian result in the first approximation. Let us now compare Birkhoff's and Einstein's theories.

Birkhoff's theory:

Flat space-time: $ds^2 = dt^2 - dx^2 - dy^2 - dz^2$.

Gravitational tensor: $h_{ij} = \begin{Vmatrix} m/r & 0 & 0 & 0 \\ 0 & m/r & 0 & 0 \\ 0 & 0 & m/r & 0 \\ 0 & 0 & 0 & m/r \end{Vmatrix}$ with $\square h_{ij} = 0$.

Equations of motion of a particle in the plane $z = 0$

$$x'' = -\frac{mx}{r^3} - \frac{2mx}{r^3}(x'^2 + y'^2) + \frac{mx'r'}{r^2},$$

$$y'' = -\frac{my}{r^3} - \frac{2my}{r^3}(x'^2 + y'^2) + \frac{my'r'}{r^2}.$$

Advance of perihelion: $\dfrac{6\pi m}{a(1 - e^2)}$.

Bending of light rays: $\dfrac{4m}{p}$.

Red shift[5] $\Delta\lambda/\lambda$: $\frac{m}{r}$.

Degenerate form of Einstein's theory:

Curved space-time: $ds^2 = (1 - 2m/r)dt^2 - (1 + b/r)(dx^2 + dy^2 + dz^2)$.

Gravitational tensor: $H_{ij} = \begin{Vmatrix} -2m/r & 0 & 0 & 0 \\ 0 & b/r & 0 & 0 \\ 0 & 0 & b/r & 0 \\ 0 & 0 & 0 & b/r \end{Vmatrix}$.

Equations of motion of a particle in the plane $z = 0$

$$x'' \doteq -\frac{mx}{r^3} - \frac{bx'r'}{r^2} - \frac{(m - b/2)x}{r^3}(x'^2 + y'^2),$$

$$y'' \doteq -\frac{my}{r^3} - \frac{by'r'}{r^2} - \frac{(m - b/2)y}{r^3}(x'^2 + y'^2).$$

Advance of perihelion: $\frac{-3b + 2m}{a(1 - e^2)}\pi$.

Bending of light rays: $\frac{2m - b}{p}$.

Red shift $\Delta\lambda/\lambda$: $\frac{m}{r}$.

From this it follows that, in order to obtain the correct advance of the perihelion one must place $b = -4m/3$ and, in order to obtain the correct bending of light, one should place $b = -2m$ which is contradictory.

The red shift is accounted for in Birkhoff's theory as due to the energy loss of the photon while it travels from the emitting body to the earth. On the basis of this alternative interpretation, but using the degenerate form of the equations, b would then have to vanish, $b = 0$, to obtain the correct result.

On the other hand, in the *flat* space-time of the degenerate form of Einstein's theory the atomic clock would register imaginary time at distances less than $2m$ from the mass-point at the origin!

The introduction of the two arbitrary constants a and b is not in keeping with the spirit of Einstein's original theory, one of whose chief advantages is the elimination of arbitrary constants. Besides, if in the energy tensor $T^{ij} = \rho u^i u^j - pg^{ij}$ one sets $p = 0$, in accordance with the fact that the potential pressure energy is small, as Einstein has done, then $b = 0$ necessarily follows.

Another characteristic feature of Birkhoff's theory should be carefully emphasized: the trajectories of free particles are not geodesics in any curved four-dimensional space-time, i.e., no $d\sigma^2 = g_{ij}dx^i dx^j x$ exists in which Birkhoff's trajectories are geodesics. In fact for light rays one would have, on the one hand $d\sigma^2 = g_{ij}dx^i dx^j = 0$ and on the other $ds^2 = \delta_{ij}dx^i dx^j = 0$, from which $g_{ij} = k\delta_{ij}$, k = constant, would follow. Otherwise, we may establish our statement thus: the equations of the geodesics in the $d\sigma$-space are:

$$u^p\left(\frac{d^2x^i}{ds^2} + \Gamma^i_{\alpha\beta}u^\alpha u^\beta\right) = u^i\left(\frac{d^2x^p}{ds^2} + \Gamma^p_{\alpha\beta}u^\alpha u^\beta\right)$$

Multiplying by u_i and adding with respect to i one obtains:

$$u^p u^i \Gamma^i_{\alpha\beta}u^\alpha u^\beta = \frac{d^2x^p}{ds^2} + \Gamma^p_{\alpha\beta}\, u^\alpha u^\beta$$

from which we get

$$\frac{d^2x^p}{ds^2} = (u^p u_i \Gamma^i_{\alpha\beta} - \Gamma^p_{\alpha\beta})\, u^\alpha u^\beta.$$

Now the trajectories of free particles in Birkhoff's theory are given by

$$\frac{d^2x^p}{ds^2} = \delta^{\mu p}\left(\frac{\partial h_{\alpha\mu}}{\partial x^\beta} - \frac{\partial h_{\mu\beta}}{\partial x^\mu}\right)u^\alpha u^\beta;$$

therefore one should have identically[6]

$$\delta^{\mu p}\left(\frac{\partial h_{\alpha\mu}}{\partial x^\beta} - \frac{\partial h_{\alpha\beta}}{\partial x^\mu}\right) = u^p u_i\, \Gamma^i_{\alpha\beta} - \Gamma^p_{\alpha\beta}.$$

But the left-hand member of this equation is a function of t, x, y, z, only, while the right-hand member is a function of t, x, y, z, and u^k; therefore the two members cannot be identical unless $\Gamma^i_{\alpha\beta} = 0$ which is a trivial case. This completes our proof.[7]

---

[1] *Mathematical Reviews*, **4**, 285 (Nov. 1943).

[2] These PROCEEDINGS, **29**, 231 (1943).

[3] During the last two months G. D. Birkhoff, Carlos Graef, Manuel Sandoval Vallarta and I have been collaborating intensively in a study of Birkhoff's theory of gravitation, matter and electricity in flat space-time, and its consequences. The present note owes much to our joint discussions.

[4] See, for instance, R. C. Tolman, *Relativity, Thermodynamics and Cosmology*, Oxford, 1934, p. 236.

[5] Obtained by use of photon as light carrier. All formulas are taken only in the first approximation.

[6] The components of u^i are related by $\delta_{ij}u^i u^j = 1$.

[7] Weyl's other comments will be taken up separately by Manuel Sandoval Vallarta.

ON THE THEORY OF CHARACTERS OF GROUPS AND SEMI-GROUPS IN NORMED VECTOR RINGS

BY EINAR HILLE

DEPARTMENT OF MATHEMATICS, YALE UNIVERSITY

Communicated February 4, 1944

1. Recently I. Gelfand[1] has defined a class of characters of Abelian groups embedded in a normed commutative vector ring. The present note contains an elaboration of this theory with a strengthening of the main theorem and extensions to semi-groups.

Let $\mathfrak{R}$ be a normed commutative vector ring with unit element e, $\boldsymbol{\mathfrak{M}}$ the set of maximal ideals $\mathfrak{M}$ in $\mathfrak{R}$, $\mathfrak{R}/\mathfrak{M}$ the corresponding residue class rings each of which is a field. Defining the norm $||X||$ of a residue class X, modulo $\mathfrak{M}$, as min $||x||$ for $x \epsilon X$, Gelfand shows that $\mathfrak{R}/\mathfrak{M}$ is a normed field and hence isomorphic with the complex field. Let X_α be the class of elements congruent to αe (mod. $\mathfrak{M}$), α arbitrary complex number. Then $\bigcup X_\alpha = \mathfrak{R}$ and $X_\alpha + X_\beta = X_{\alpha+\beta}$, $X_\alpha X_\beta = X_{\alpha\beta}$.

Define[2]

$$\mu(x;\ \mathfrak{M}) = \alpha, \qquad x \equiv \alpha e \text{ (mod. } \mathfrak{M}). \tag{1}$$

Since $x - \alpha e$ is singular for every $x \epsilon X_\alpha$, the spectrum $\sigma(x)$ of x contains the point $\lambda = \alpha$. Conversely, if $\alpha \epsilon \sigma(x)$, there exists an $\mathfrak{M}$ such that $x \equiv \alpha e$ (mod. $\mathfrak{M}$). Since $|\mu(x;\ \mathfrak{M})| \leqq \max |\sigma(x)| \leqq ||x||$, $\mu(x;\ \mathfrak{M})$ is a bounded functional which is additive as well as multiplicative

$$\mu(x + y;\ \mathfrak{M}) = \mu(x;\ \mathfrak{M}) + \mu(y;\ \mathfrak{M}), \mu(xy;\ \mathfrak{M}) = \mu(x;\ \mathfrak{M})\mu(y;\ \mathfrak{M}). \tag{2}$$

In particular, $\mu(x;\ \mathfrak{M})$ is continuous.

Finally the notion of quasi-nilpotency will be needed. An element q of $\mathfrak{R}$ is quasi-nilpotent if $e - \alpha q$ has an inverse for every complex α or, which is equivalent in a normed vector ring, if $||q^n||^{1/n} \to 0$ when $n \to \infty$. The quasi-nilpotent elements form an ideal, the radical of $\mathfrak{R}$, which equals the intersection of all maximal ideals.

After these introductory remarks we proceed to the subject matter proper.

2. Let $\mathfrak{G}$ be a set of elements in the ring $\mathfrak{R}$ which form a group under the operation of ring multiplication. For every $\mathfrak{M} \epsilon \boldsymbol{\mathfrak{M}}$ the bounded linear multiplicative functional $\mu(x;\ \mathfrak{M})$ for x in $\mathfrak{G}$ defines a continuous one-dimensional representation $\Gamma(\mathfrak{M})$ of $\mathfrak{G}$, $\mu(x;\ \mathfrak{M})$ being the characters of the representation.

THEOREM 1. *$\Gamma(\mathfrak{M})$ is unitary in the sense that $|\mu(x;\ \mathfrak{M})| = 1$ on $\mathfrak{G}$ for every $\mathfrak{M} \epsilon \boldsymbol{\mathfrak{M}}$ if and only if for all $x \epsilon \mathfrak{G}$*

$$\lim_{n\to\infty} ||x^n||^{1/n} = 1, \qquad \lim_{n\to\infty} ||x^{-n}||^{1/n} = 1. \tag{3}$$

Condition (3) expresses that the spectrum of x lies on the unit-circle. Since $\sigma(x)$ is the range of $\mu(x;\ \mathfrak{M})$ for $\mathfrak{M} \in \boldsymbol{\mathfrak{M}}$ and fixed x, this is all that should be proved.

The main question in the representation theory is whether or not there are enough representations available so that any given pair of distinct elements of the group have distinct characters in at least one representation. The limitations of the present theory are clearly brought out by

THEOREM 2. *If* $x \neq y$ *are two elements of* $\mathfrak{G}$, *then* $\mu(x;\ \mathfrak{M}) = \mu(y;\ \mathfrak{M})$ *for all* $\mathfrak{M} \in \boldsymbol{\mathfrak{M}}$ *if and only if* $x - y = q$, *where* q *is quasi-nilpotent. This case can arise if and only if* $\mathfrak{G}$ *contains an element of the form* $e + q$ *and is excluded if* $\mathfrak{R}$ *is without radical.*

We omit the proof which is an immediate consequence of the properties of the functional $\mu(x;\ \mathfrak{M})$ listed above.

Gelfand has shown that even if $\mathfrak{R}$ has a radical it is possible to exclude the exceptional elements of the form $e + q$ from the group $\mathfrak{G}$ by requiring that all cyclical subgroups form bounded pointsets in the vector space $\mathfrak{R}$. The following condition is less restrictive.

THEOREM 3. *If* $||x^{\pm n}|| = o(n)$ *when* $n \to \infty$, *then* x *cannot be of the form* $e + q$ *where* q *is quasi-nilpotent and* $q \neq 0$. *The conclusion becomes false if* $o(n)$ *be replaced by* $O(n)$.

If $x = e + q$ then the resolvent $R(\lambda;\ e + q) = ((\lambda - 1)e - q)^{-1}$ admits of the three expansions

$$\begin{aligned} -(e+q)^{-1} - \textstyle\sum_1^\infty (e+q)^{-n-1}\lambda^n, &\qquad |\lambda| < 1; \\ e\lambda^{-1} + \textstyle\sum_1^\infty (e+q)^n\lambda^{-n-1}, &\qquad |\lambda| > 1; \\ e(\lambda - 1)^{-1} + \textstyle\sum_1^\infty q^n(\lambda - 1)^{-n-1}, &\qquad \lambda \neq 1. \end{aligned}$$

In the first two series the coefficients are $o(n)$ in norm. A theorem in classical function theory, found by G. Pólya,[3] asserts that if $f(z)$ is an entire function of $w = 1/(z - 1)$ and if in the power series expansions of $f(z)$ about 0 and about ∞ the coefficients are $O(n^\nu)$, then $f(z)$ is a polynomial in w of degree $\leqq \nu + 1$. If the coefficients are $o(n^\nu)$ instead, the degree will be $< \nu + 1$. By an obvious extension of this theorem to vector valued functions, one concludes that $R(\lambda;\ e + q)$ has a simple pole at $\lambda = 1$ which requires $q = 0$. On the other hand, if q is a nilpotent element such that $q \neq 0$, $q^2 = 0$, then $(e + q)^n = e + nq$ which is $O(n)$ in norm. Thus the stated condition is the best of its kind.

It should be observed that a condition of the form $||x^{\pm n}|| = o(n^m)$, if satisfied by an element of the form $e + q$, requires $q^m = 0$ and this condition is again the best possible.[4]

Combining the preceding theorems we obtain

THEOREM 4. *If the elements of $\mathfrak{G}$ satisfy the condition of Theorem 3, then all representations* $\Gamma(\mathfrak{M})$ *are unitary and if* $x \neq y$, *there is a representation in which* $\mu(x;\ \mathfrak{M}) \neq \mu(y;\ \mathfrak{M})$.

3. These considerations may be extended in part to semi-groups. Let $\mathfrak{S}$ be a set of elements in $\mathfrak{R}$ which form a semi-group under ring multiplication so that $x \in \mathfrak{S}$, $y \in \mathfrak{S}$ implies $xy \in \mathfrak{S}$. We shall require that $\mathfrak{S}$ be closed under inversion so that if x is regular and is in $\mathfrak{S}$ then x^{-1} is also in $\mathfrak{S}$.

The functionals $\mu(x;\ \mathfrak{M})$ also provide one-dimensional representations $\Sigma(\mathfrak{M})$ of $\mathfrak{S}$. It is obviously meaningless to require that $\Sigma(\mathfrak{M})$ be unitary. If $x \neq y$ are two elements of $\mathfrak{S}$ we still have $\mu(x;\ \mathfrak{M}) = \mu(y;\ \mathfrak{M})$ for all $\mathfrak{M}$ if and only if $x - y = q$. Thus this case can arise only if $\mathfrak{R}$ has a radical, but when present the situation is much more complicated than in the case of groups. Thus as a rule there is no longer a single exceptional class; every singular element $x \in \mathfrak{S}$ may give rise to a class $\{x + q\}$ the elements of which cannot be separated by the representations. Since the condition of Theorem 3 can be applied to regular elements only, the problem of excluding such singular exceptional classes is a much more difficult one which apparently cannot be handled by the methods of the present note.

---

[1] Gelfand, I., "Zur Theorie der Charaktere der Abelschen topologischen Gruppen," *Rec. Math.*, N. S., **9** (**51**), 49–50 (1941). For the theory of normed vector rings see Gelfand, "Normierte Ringe," *Ibid.*, 3–24, and Lorch, E. R., "The Theory of Analytic Functions in Normed Abelian Vector Rings," *Trans. Amer. Math. Soc.*, **54**, 414–425 (1943).

[2] This is $x(\mathfrak{M})$ in Gelfand's notation. His $x(\mathfrak{M})$, however, is a function on maximal ideals to complex numbers corresponding to a fixed element x of the ring while $\mu(x;\ \mathfrak{M})$ is a function on the ring to complex numbers corresponding to a fixed maximal ideal.

[3] Pólya, G., "Aufgabe 106," *Jahresbericht D.M.V.*, **40**, *81* (1931). Several solutions are published, *Ibid.*, **42** (1933); that by Szegö, G., pp. *15–16*, is short and self-contained. See also the closely related problem 105: an entire function of order one and minimal type which is $O(n^m)$ on the integers is a polynomial of degree $\leq m$.

[4] Cf. Gelfand, I., "Ideale und primäre Ideale in normierten Ringen," *Rec. Math.*, N. S., **9** (**51**), 41–48 (1941), where similar conclusions are drawn less directly from a deeper theorem due to F. and R. Nevanlinna.

GALACTIC AND EXTRAGALACTIC STUDIES, XVII. REVISION OF THE DISTANCES OF THIRTY HIGH-LATITUDE GLOBULAR CLUSTERS

BY HARLOW SHAPLEY

HARVARD COLLEGE OBSERVATORY

Communicated February 7, 1944

1. The distribution of the photoelectrically-measured color classes for sixteen globular clusters in galactic latitudes higher than thirty degrees is as follows:[1]

Color class	*a*5	*f*4	*f*5	*f*6	*f*7	*f*8	*f*9
Numbers	1	1	1	5	1	6	1

with the mean value *f*6.1 ± 0.8(m. e.). The distribution of integrated spectral types, according to Miss Cannon's classification for ten globular clusters in these same latitudes is:

Spectral class	*A*7	*F*5	*F*8	*G*0	*G*	*G*5
Numbers	1	1	1	1	3	3

The difficulty of classification is such that only the mean value of about *G*0 is significant. For the seven high-latitude globular clusters common to both lists we have the mean values, *f*6.7 and *F*8. The clusters certainly are not redder than would be expected from their mean spectral classes.

The foregoing evidence that there is little or no selective space absorption in high latitude is supported also by the work of Stebbins, Whitford and Huffer on the color excesses for stars and galaxies. The distribution of faint galaxies indicates that also the total photographic absorption in high latitude, if not wholly negligible, is at least manageably small. Even down to latitudes of twenty degrees the total absorption, except in a few special longitudes, is not so heavy but that we can, when computing the distances of globular clusters, estimate the amount of light loss from the abundance of external galaxies in the background.

For the sixty known globular clusters of the galactic system in low latitude, $\beta < \pm 20°$, accurate determinations of the distance will long remain difficult. For them the background of external galaxies is absent, or nearly so, and the total absorption cannot therefore be directly estimated. It is precarious to adopt an absorption correction from the observed color excesses because of the large inherent spread in the integrated colors and the possible variability of the ratio of selective to total absorption throughout the various longitudes. All the photometric methods of measuring either the absolute or the relative distances are therefore not encouraging for clusters in low latitude, even when many

Cepheid variables are present; and the astrometric methods (parallax, motions, diameters) are of low value, either because the distances are too great or because of vagueness in the boundaries of clusters in the rich star fields.

Although occasionally in intermediate latitudes, $\pm 20^\circ > \beta > \pm 10^\circ$, the total absorption in the direction of a globular cluster can be estimated from the number of external galaxies in the surrounding field, we may be at a loss to know how much absorption lies between the cluster and the observer and how much lies beyond the cluster but still within the galactic system. All high-latitude globular clusters, however, lie well beyond whatever absorption there may be near the galactic plane, and for them the frequency of surrounding external galaxies is a fair measure of the intervening space absorption in the cluster's direction.

New information on interstellar transparency is now available for much of the sky from the surveys of faint galaxies, and we can proceed to revise the distances of globular clusters that are not too near the galactic plane. For many of them we now have revised distance moduli, based on new photometry of the variables or the bright cluster stars from the David Dunlap, Harvard, Mount Wilson and other observatories. In the far southern sky, where many of the high-latitude clusters are found, the photographic magnitude system, for stars as faint as those in many of the globular clusters, is still provisional.

2. In table 1 we have assembled, chiefly from *Star Clusters*,[2] some descriptive data concerning the globular clusters for which revised values of the distances have now been determined. Galactic coördinates λ and β, refer to the pole at 12^h40^m, $+28^\circ$ (1900). The angular diameters, except one in parentheses, are those microdensitometrically measured by Shapley and Sayer.[3] The integrated magnitudes are, when possible, taken from Christie's schraffiekassette work;[4] otherwise they are values determined from Harvard patrol plates and are reduced to the Christie scale.

3. The new results for the high-latitude globular clusters are collected in table 2. In the first section of the table are the objects for which the present results are the most dependable. The lower reliability for the clusters in the second section arises from various sources; in general, the observed quantities that contribute to the low weight are indicated by colons. The best one-half of the entries in the second section are marked by asterisks after the NGC number. A letter *n* refers to a note at the end of the table.

The third, fourth and fifth columns of table 2 give the results of the nebular counts on the best available long-exposure plates made with the Bruce and Metcalf telescopes for the southern and northern skies, respectively. The average number of galaxies per square degree, $\overline{N}$, generally refers to a surrounding or adjacent area of not less than nine square de-

TABLE 1

NGC	R. A. (1900)	DEC.	λ	β	DIAMETER	TOTAL PG. MAGNITUDE	NO. OF VARIABLES
104	$0^h19^m.6$	−72°38′	272°	−45°	53′.6	(4.5)	8*
288	0 47.8	−27 08	157	−88	13.1	8.96	2*
362	0 58.9	−71 23	268	−47	17.7	8.0	14
1261	3 9.5	−55 36	237	−51.5	...	9.5	0
1851	5 10.8	−40 09	212	−34.5	11.5	7.72	3*
2419	7 31.4	+39 06	148	+26	5.1	11.51	36
4147	12 5.0	+19 06	226	+79	2.8	11.01	4
4590	12 34.2	−26 12	269	+36	...	9.12	28
5024	13 8.0	+18 42	305	+79	15.8	8.68	40
5053	13 11.5	+18 13	310	+78	...	10.9	9
5272	13 37.6	+28 53	8	+78	22.1	7.21	185
5466	14 1.0	+29 00	8	+72.5	...	10.39	14
5634	14 24.4	− 5 32	310	+48.5	6.8	10.8	4*
5694	14 33.8	−26 36	299	+29	...	10.87:	0
5897	15 11.7	−20 39	312	+29	13.1	9.61	0
5904	15 13.5	+ 2 27	332	+46	25.0	7.04	97
6205	16 38.1	+36 39	27	+40	18.1	6.78	11
6218	16 42.0	− 1 46	344	+25	21.0	7.95	1
6229	16 44.2	+47 42	40	+40	5.3	10.26	21*
6254	16 51.9	− 3 57	343	+22	21.5	7.64	2
6341	17 14.1	+43 15	36	+35	(14)†	7.30	16
6752	19 2.0	−60 48	303	−26.5	41.9	7.2:	1*
6809	19 33.7	−31 10	336	−25	28.7	7.08	2*
6864	20 0.2	−22 12	347	−27	8.7	9.50	11*
6934	20 29.3	+ 7 04	20	−20	...	10.01	51*
6981	20 48.0	−12 55	3	−34	8.0	10.24	31
7006	20 56.8	+15 48	32	−21	...	11.45	20
7078	21 25.2	+11 44	33	−28	18.1	7.33	74
7089	21 28.3	− 1 16	21	−36	16.9	7.30	17
7492	23 3.1	−16 10	22	−64	...	12.33	9*

* Distances do not depend on variable stars.

† Diameter of NGC 6341 from Nassau, J. J., *Ap. J.*, **87**, 366 (1938).

grees. The magnitude limit to which the count of galaxies refers is m_n, and the absorption δm is computed from the relation

$$\delta m = m_n - 1.67 \log \overline{N} - 15.2$$

which is based on our earlier determinations[5] of the space-density parameter for the galactic caps, $\overline{m}_1 = 15.2$.

The number of Cepheid variable stars that have been recorded in each cluster is shown in the last column of table 1.[6] When it was not possible to use the variables directly in the determination of the distance modulus, the number is marked with an asterisk; the modulus has then been determined from the cluster's bright stars and the integrated magnitude and diameter, which is also the procedure when no variables at all are known. The adopted distance modulus in the sixth column of table 2 includes the

TABLE 2a

NGC	MESSIER	$\ddot{\bar{N}}$	m_n	δm	DIST. MOD.	DIST. KPC	$r \sin \beta$ KPC	$r \cos \beta$ KPC	ABS. MAG.
104	47 Tuc	*n*	*n*	+0.3	14.4	7.6	− 5.4	5.4	−10.2
288		85	18.4	0.0	15.8	14.5	−14.5	10.5	−6.8
362		*n*	*n*	+0.3	15.0	10.0	− 7.3	6.8	−7.3
2419		*n*	*n*	+0.5	18.75	56.2	+24.6	50.4	−7.7
4147		76*n*	17.5	−0.8:	16.5	20.0	+19.6	3.8	−5.5
4590	68	8	17.0	+0.3	15.65	13.5	+ 7.9	10.9	−6.8
5024	53	41	17.7	−0.2	16.53	20.2	+19.8	3.9	−7.8
5053		41	17.7	−0.2	16.2	17.4*n*	+17.0	3.6	−5.3
5272	3	32*n*	17.2	−0.5	15.43	12.2	+11.9	2.6	−8.2
5466		56*n*	17.5	−0.6:	16.16	17.0	+16.2	5.1	−5.8
5694		*n*	*n*	+0.4:	17.6	33.1	+16	29	−7.1:
5897		19	17.7	+0.4	15.7	13.8	+ 6.7	12.1	−6.5
5904	5	26	17.6	0.0	15.03	10.1*n*	+ 7.3	7.0	−8.0
6205	13	14	17.0	−0.1	14.88	9.5*n*	+ 6.1	7.3	−8.1
6218	12	8	17.3	+0.6	14.6	8.3	+ 3.5	7.5	−7.3
6254	10	8	17.3	+0.6	14.6	8.3	+ 3.1	7.7	−7.6
6341	92	34	17.4	−0.3	15.06	10.3	+ 5.9	8.4	−7.8
6981	72	12	17.7	+0.7	16.1	16.6	− 9.3	13.8	−6.6
7078	15	44	18.2	+0.3	15.3	11.5	− 5.4	10.2	−8.3
7089	2	33	17.8	+0.1	15.7	13.8	− 8.1	11.2	−8.5

TABLE 2b

NGC	MESSIER	$\ddot{\bar{N}}$	m_n	δm	DIST. MOD.	DIST. KPC	$r \sin \beta$ KPC	$r \cos \beta$ KPC	ABS. MAG.
1261		21	17.3	−0.1	16.7:	22	−17	14	−7.2
1851		68	17.7	−0.5	15.8:	14	− 8	12	−8.1
5634		16	17.2	0.0	17.5:	32	+24	21	−6.7
6229*		27*n*	17.1	−0.5:	17.4	30*n*	+19	23	−7.1
6752		17	18.0	+0.8	13.8:	5.8	− 2.6	5.2	−7.4:
6809*	55	3	16.9	+0.9	13.8	5.8	− 2.4	5.3	−7.7
6864	75	8	17.0:	+0.3:	18.1:	42	−19	37	−8.9:
6934*		24:	18.2	+0.7:	16.3	18	− 6	17	−7.0
7006*		19	17.9	+0.6	18.2:	44	−16	41	−7.3
7492*		33	18.2	+0.5	17.0	25.1	−23	11	−4.7

NGC 104. See twelfth paper of this series, these PROCEEDINGS **27**, 440–445 (1941), for discussion of distance and absorption.

362. Absorption assumed the same as for the Small Magellanic Cloud, which surrounds it.

2419. Adopted value of absorption from Baade's discussion in *Mt. Wilson Contr.*, No. 529 (1935). A Harvard plate suggests a larger value, but the revision awaits further study of the surrounding regions.

4147, 5272, 5466, 6229. Clusters lie in regions rich in groups of galaxies.

5053. By comparing the color-magnitude array for this cluster with those I derived for Messier 3, 13 and 68, Cuffey has estimated the distance as 16 ± 2 kpc. *Ap. J.*, **98**, 49–53 (1943). This cluster and NGC 4147, 5466 and 7492 belong to the so-called "giant-poor" category, with low absolute magnitudes.

5694. Baade reports little or no absorption (*Pub. Ast. Soc. Pac.*, **46**, 52 (1934)); my provisional estimate is $\delta m = +0.8$ from Bruce plates.

5904. The revised distance is based on my earlier general photometry and on a new study of the variables by Oosterhoff at Leiden (letter, dated December 19, 1940).

6205. Distance of Messier 13 revised with assistance of H. B. Sawyer's work on four Cepheids. *Pub. David Dunlap Obs.*, 1, 11 (1942).

6229. Distance can be precisely determined when the 20 new variables (unpublished), discovered by Baade at Mt. Wilson, are worked up.

correction for space absorption, column 5, and involves also whatever revision we have been able to make on the basis of recent photometric work at Harvard, Mount Wilson (Baade), Perkins (Nassau), David Dunlap (Sawyer), Babelsberg (Hachenberg), Leiden (Oosterhoff) and Bergedorf (Grosse) observatories.

No absorption correction has been applied when δm is negative. When it is positive and the latitude is greater than ±40°, also no correction has been made, except for the two clusters NGC 104 and 362 which are near the Small Magellanic Cloud and appear to suffer from its absorbing dust or from a flare of absorption extending from the Milky Way. The negative values of δm reflect chiefly the uneven distribution and in particular the excesses in the number of faint galaxies; we must, of course, expect that the positive values frequently reflect true deficiencies in number rather than space absorption. Figure 1 shows the distribution of δm.

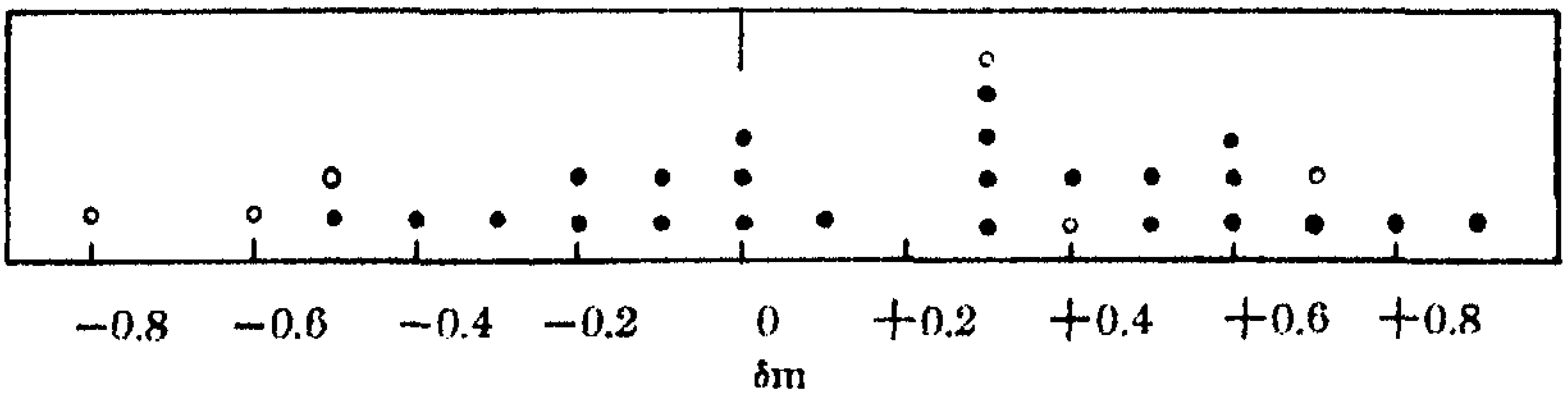


FIGURE 1

Distribution of absorption corrections. Low-weight values indicated by circles.

The distances in the seventh to ninth columns of table 2 are "reduced to the galactic poles"; that is, they are calculated on the assumption that there is no absorption in the galactic polar caps—an assumption which implies that the Sun is located in an open space in the scattered and irregular clouds of absorption along the Milky Way plane. The components of the distance perpendicular to and in the galactic plane are given in the eighth and ninth columns, respectively; β is the galactic latitude.

As mentioned above, the evidence from the nebular distribution and from the colors of clusters and galaxies suggests that the assumption of polar-cap transparency cannot be far wrong. If there is an appreciable absorption at the galactic poles, all the distances in columns seven, eight and nine of table 2 should be multiplied by the following factors:

If total absorption at pole is $0^{\rm m}.1$, $f = 0.955$
0.2 0.91
0.3 0.87

Hubble has used $0^{\rm m}.25$ for the total photographic absorption at the galactic poles; Stebbins and Whitford, a somewhat smaller value.

If such a correction factor should be applied to the distances in table 2, it probably would be appropriate first to use all the positive and at least the smaller negative corrections for absorption, rather than follow the procedure described above. Such a plan would merely imply that we adopt a specific *average* space absorption for high galactic latitudes, but that here and there we find either excessive absorption and a positive δm, or clear spaces with consequent high nebular number and a negative absorption correction.

TABLE 3

COMPARISON OF REVISED DISTANCES WITH 1930 VALUES

NGC	GALACTIC LATITUDE	1930 VALUE KPC	RATIO	NGC	GALACTIC LATITUDE	1930 VALUE KPC	RATIO
104	−45°	6.8	1.12	6205	+40°	10.3	0.93
288	−88	14.5	1.00	6218	+25	11.0	0.75
362	−47	12.9	0.78	6229	+40	29.8	1.01
1261	−51.5	22.0	1.00	6254	+22	11.2	0.74
1851	−34.5	14.3	0.98	6341	+35	11.2	0.92
4147	+79	24.2	0.83	6752	−26.5	8.4	0.69
4590	+36	15.5	0.87	6809	−25	8.8	0.66
5024	+79	18.2	1.11	6864	−27	48.5	0.87
5053	+78	17.3	1.00	6934	−20	24.9	0.72
5272	+78	12.2	1.00	6981	−34	23.3	0.71
5466	+72.5	17.0	1.00	7006	−21	56.8	0.78
5634	+48.5	31.4	1.02	7078	−28	13.1	0.88
5897	+29	16.4	0.84	7089	−36	13.9	0.99
5904	+46	10.8	0.94	7492	−64	25.2	0.99

4. The new results for the distances are compared in table 3 with my original values, published on the currently-accepted zero point in 1930.[2] Naturally there is on the average little change for the objects in the higher latitudes; we have the mean ratio New/Old = 0.98 ± 0.03 (m. e.) for the twelve clusters with latitudes higher than 40°. The change is considerable, however, with latitudes from 20° to 40°, where for sixteen clusters we have, in the mean, New/Old = 0.83 ± 0.03 (m. e.). The corrections for space absorption produce most of the change.[7]

The distances from the galactic plane, $r \sin \beta$, in column 8 of table 2, show results now well known—namely, the great thickness of the Milky Way system, if we assume, as is natural, that most if not all the high-latitude globular clusters are members of the Galaxy. It is of interest to compare the largest values, around 20 kiloparsecs above and below the galactic plane with the values r sin β for the Magellanic Clouds, 12 kiloparsecs (Large) and 18 kiloparsecs,[8] and for the faintest of the independent high-latitude cluster-type Cepheids, 10 to 15 kiloparsecs.[9]

5. In the final column of table 2 are the absolute photographic magnitudes of the thirty globular clusters. Their dispersion is shown in figure 2, which resembles a diagram given by Christie.[4]

The greatest of these globular clusters, 47 Tucanae, exceeds in total absolute brightness some of the dwarf irregular galaxies of the Magellanic type. The globular cluster Omega Centauri is of comparable absolute magnitude, but its low latitude, $\beta = +15°$, excludes it from the tabulations above. For it a revised photometry by Martin,[10] and an unpublished Harvard nebular survey that suggests an absorption of +0.5, would give it the distance of 6.8 kiloparsecs, the same as my 1930 value, and an absolute magnitude of −10.0.

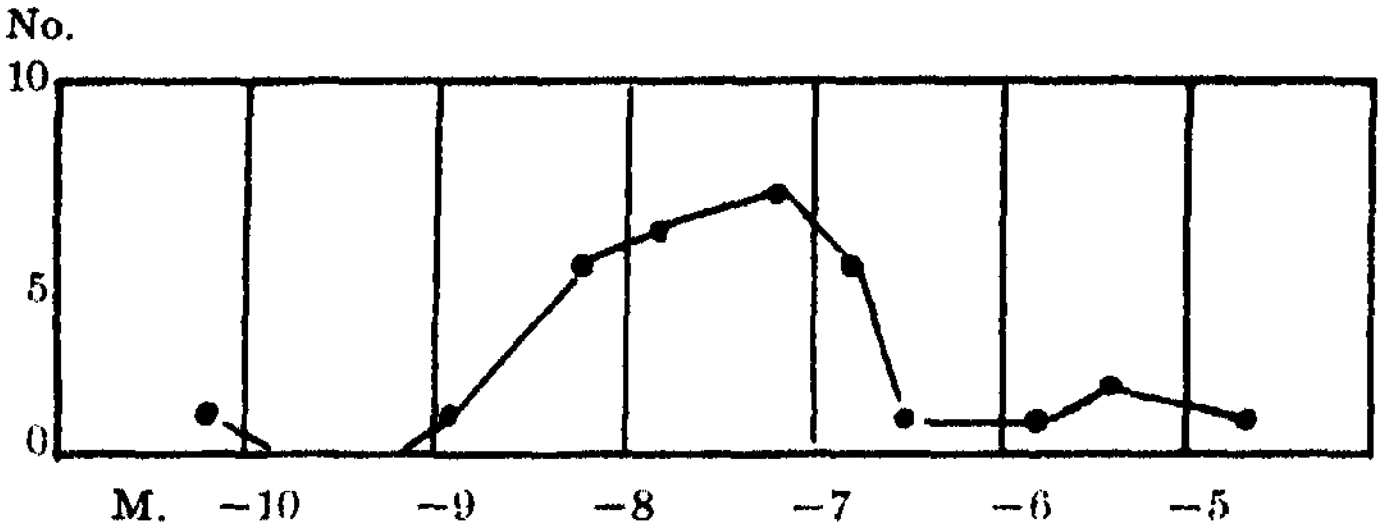


FIGURE 2

Thirty globular clusters. Spread of absolute magnitude.

6. *Conclusions and Summary.*—(*a*) The distances of thirty globular clusters in intermediate and high galactic latitudes have been revised on the basis of (1) recent investigations of the photographic magnitudes of bright stars and of Cepheid variables in the clusters, and (2) a census of the background of external galaxies, which gives evidence on the transparency of space between the observer and the clusters.

(*b*) The over-all diameter of the Milky Way in its plane is not well indicated by the revised distances, because of the necessary avoidance of low-latitude objects; but the earlier estimates of a distance to the center of approximately ten kiloparsecs are consistent with the new values of the distances of clusters.

(*c*) The over-all thickness of the system of globular clusters exceeds forty kiloparsecs. The weighted mean r sin β for the five clusters most distant from the galactic plane on the north side is +19.3 kiloparsecs, and for the five most distant on the south side, −17.5 kiloparsecs. The clusters apparently have a spatial distribution comparable with the haze of cluster-type Cepheids discussed in the first paper of this series.

(*d*) The integrated absolute magnitudes of the thirty globular clusters show a small dispersion around the mean value of -7.5 (photographic), except that the great globular cluster 47 Tucanae stands out as nearly three magnitudes brighter than the mean (as does Omega Centauri), and the "giant-poor" clusters, NGC 4147, 5053, 5466, 7492, are all fainter than the average, with values lying between -4.8 and -5.9. It is noteworthy that the absolute magnitudes of several of the dwarf galaxies of the local group of galaxies are between -9 and -12; and, of probably much higher significance in the hypothetical evolution of galaxies, the absolute magnitudes of the nuclei of the spirals NGC 598 (Messier 33) and 7793 are -8.0 and -7.3, respectively. The "escape" of the spiral arms and their background of stars, in a developmental trend from spiral to globular form,[11] would leave these two objects as globular clusters of average absolute luminosity.

[1] Stebbins and Whitford, *Mt. Wilson Contr.*, No. 547, p. 10 f (1936).

[2] *Harv. Obs. Monographs*, No. 2, Appendix A (1930).

[3] These PROCEEDINGS, **21**, 593–597 (1935).

[4] *Mt. Wilson Contr.*, No. 620 (1940).

[5] These PROCEEDINGS, **26**, 599–604 (1940), the tenth paper of this series.

[6] From Helen B. Sawyer's compilation in *Pub. David Dunlap Obs.*, **1**, No. 4 (1939), with some additions.

[7] Two clusters are omitted from the comparison: NGC 2419, the "intergalactic tramp," and NGC 5694 which was identified as a globular cluster after the earlier study: Lampland and Tombaugh, *Ast. Nach.*, **246**, 171 (1932).

[8] *Galaxies*, Blakiston, 1943, p. 51.

[9] These PROCEEDINGS, **25**, 423–428 (1939), the first paper of this series.

[10] *Leiden Obs. Ann.*, **17**, pt. 2, 1–166 (1938).

[11] *Galaxies*, Blakiston, 1943, p. 216.

PROCEEDINGS
OF THE
NATIONAL ACADEMY OF SCIENCES

Volume 30 April 15, 1944 Number 4

Copyright 1944 by the National Academy of Sciences

THE BEHAVIOR OF OPPOSITIONAL ALLELES IN POLYPLOIDS OF TRIFOLIUM REPENS[1]

BY SANFORD S. ATWOOD

U. S. REGIONAL PASTURE RESEARCH LABORATORY, STATE COLLEGE, PA.

Communicated February 16, 1944

The theory of oppositional alleles was first proposed with substantiating data by East and Mangelsdorf[2] to explain self- and cross-incompatibilities in *Nicotiana*, and subsequently it was confirmed by other workers in several diploid species. Corollary hypotheses have been proposed to extend its application to polyploids, but because of the paucity of experimental data no hypothesis has been verified completely. The colchicine technique, which allows production of polyploids from plants of known genotype, should provide a more precise approach to this problem, but the results published so far have been conflicting. Stout and Chandler[3] reported that 15 diploid plants of *Petunia axillaris* were self-incompatible whereas tetraploid branches induced on the same plants were self-compatible. The same authors[4] showed that the 212 seedlings obtained from the tetraploid branches were likewise self-compatible and within families were cross-compatible. A similar situation was described by Crane and Lewis[5] for pears wherein most of the 29 varieties tested were self-incompatible or of low self-compatibility. The exceptions were "Fertility 2 X," which sets fruits but no seeds, and "Fertility 4 X," which sets both fruits and seeds. In contrast with these results, Howard[6] reported that both types of branches on diploid-tetraploid sectorial chimeras in *Brassica rapa*, *B. campestris* and *Raphanus sativus* have always been self-incompatible. Likewise Lewis[7] found the autotetraploid branches of *Oenothera organensis* to be self-incompatible like the diploid; the inhibition of pollen tubes, which was strong in the diploid, was reduced in the tetraploid but not enough to allow fertilization. In extensive tests of autotetraploid maize, Randolph[8] and Fischer[9] observed that different stocks varied from low to high self-fertility and that these differences were accounted for only in part by cytological irregularities. Certain self-sterile stocks were ordinarily cross-sterile as females but cross-fertile as males. There was also some evidence of cross-incompatibilities between certain self-fertile stocks. Since genes for incompatibility have not been described for diploid maize, Fischer[9] suggested that chromosome doubling results in a change in genic balance due to the cumulative action of some genes and the non-cumulative action of others.

Because suitable terms to describe multiple alleles in autotetraploids would facilitate the discussion, the following new terminology is suggested and will be used in this paper:

NO. OF DIFFERENT ALLELES	EXAMPLE	SUGGESTED TERM		DESIGNATION IN TABLE I
None	$S_1S_1S_1S_1$	Monoallelic		...
Two	$S_1S_1S_1S_2$	Unbalanced	Diallelic	...
Two	$S_1S_1S_2S_2$	Balanced		II
Three	$S_1S_1S_2S_3$	Triallelic		III
Four	$S_1S_2S_3S_4$	Tetraallelic		IV

As a result of treating white clover (*Trifolium repens* L.) seedlings with colchicine,[10] several plants were obtained on which some of the stolons had the doubled number of 64 chromosomes, whereas other stolons on the same plants retained the original number of 32. When preliminary tests of these paired sectors in the greenhouse and in the field indicated that both 32- and 64-chromosome parts were self-incompatible, six of the pairs were tested more extensively in the greenhouse. The techniques used for selfing and crossing were the same as those described previously.[11] Chromosome counts from root tips of the new cuttings showed that the 32 and 64 numbers had been maintained in each case. A rapid examination of the pollen was made with a hand lens, and every plant was found to produce a very high percentage of apparently normal, well-filled pollen. In this repeat test, both 32- and 64-chromosome cuttings again proved self-incompatible. The 32-chromosome slips yielded an average of 0.11 seed per head on the 35 heads that were self-pollinated, and the 64-chromosome slips yielded 0 seeds on 30 heads. All 61 crosses that were tried between different 32-chromosome cuttings were compatible and averaged 34.1 seeds per head. Likewise, all 31 crosses between 64-chromosome cuttings proved compatible, but they averaged only 17.5 seeds per head. The difference between these two averages was highly significant. With the 38 heads used in 64 × 32 crosses, however, only 0.29 seed per head was obtained, and with the 67 heads used in 32 × 64 crosses the average seed set was 0.51.

On the basis of these results, two 32-chromosome cuttings and the corresponding two 64-chromosome cuttings from the same two original seedlings were selected for a test of the compatibility relationships of their F_1 populations. In the 32 × 32 family, 13 F_1 plants consisted of four intra-sterile, inter-fertile groups of four, four, three and two plants, respectively, and all were reciprocally compatible with both parents. Among the 178 compatible crosses, the highest seed set per head was 56 and the lowest 23 with an average of 42.2. The differences in average seed set between the plants as females were highly significant in two of the groups (those with four plants each), but not significant in the other two. Among the same plants as males, there were no significant differences in average seed set. The maximum seed set among the incompatible crosses was 5, the minimum 0, and the average for the 35 heads (three crosses were made in duplicate) was only 0.37. The two parent plants proved self-incompatible, and the 13 F_1 plants reacted the same way. When three heads on each of the two parents were tripped to effect self-pollination, using 10 flowers per head, 0 seeds were obtained on the female and only one on the male. Likewise 0 seeds were obtained when 5 heads on each of the parents were manipulated. Similarly, the 13 F_1 plants yielded a maximum of one seed per head and an average of 0.13 when three heads per plant were tripped and an average of only 0.90 when a total of 73 heads were manipulated.

This type of self- and cross-incompatibility is explained by assuming that a single series of oppositional alleles conditions pollen-tube growth and that in this case the parents had no alleles in common:

$$S_1S_2 \times S_3S_4 = S_1S_3 + S_1S_4 + S_2S_3 + S_2S_4$$

When sectors of the 32-chromosome parent plants were doubled, the oppositional alleles likewise would have been doubled:

$$S_1S_1S_2S_2 \text{ and } S_3S_3S_4S_4$$

A cross between two such cuttings should yield the following genotypes in the F_1 generation:

DIALLELIC	TRIALLELIC	TETRAALLELIC
1 $S_1S_1S_3S_3$	4 $S_1S_1S_3S_4$	16 $S_1S_2S_3S_4$
1 $S_1S_1S_4S_4$	4 $S_2S_2S_3S_4$	
1 $S_2S_2S_3S_3$	4 $S_1S_2S_3S_3$	
1 $S_2S_2S_4S_4$	4 $S_1S_2S_4S_4$	

The frequency shown is based on random 2 × 2 assortment of the four homologous chromosomes with no crossing-over between the S gene and the centromere. With independent assortment of the eight chromatids, on the other hand, a 9:9:9:9:24:24:24:24:64 ratio would be expected, For the small population used here, it is not possible to distinguish between these two extreme ratios by means of the χ^2 test. In a study of this sort, however, the most critical evidence comes from the pattern of incompatible group reactions rather than from the close fit of obtained frequencies of different genotypes to expected ratios.

TABLE 1

Average Number of Seeds Set per Head (10 Flowers) in All Possible Self-Pollinations, Diallel Crosses and Reciprocal Backcrosses Among 29 F_1 Plants and Their Parents of 64-Chromosome White Clover. All Crosses or Selfs Included Within the Boxes Formed by the Heavy Lines Are Considered Incompatible. The Designations II, III and IV for the Different Genotypes Represent the Proposed Diallelic, Triallelic and Tetraallelic Classes

		Average[a] seeds per head											Least signif. diff. between comp. crosses and selfs
		Crosses with following ♂ genotypes:								Self-pollinations			
										Greenhouse		Field	
♀ Genotype	No. plants	II-♀ P	II-♂ P	II-1	II-2	III-1	III-2	III-3	IV	10 fls. per head	Entire heads manipulated		
II-♀ P	1	...	26.7	16.9	12.8	28.7	26.8	22.7	20.6	0.0	0.1	...	5.5
II-♂ P	1	21.0	...	20.9	16.5	23.5	22.4	21.7	23.4	0.0	0.1	...	8.5
II-1	1	7.7	$\boxed{0.0}$	...	19.5	18.5	18.3	17.7	18.2	0.0	0.7	...	6.9
II-2	2	$\boxed{0.1}$	5.7	4.3	$\boxed{0.0}$	11.0	21.0	12.8	13.6	0.2	0.0	1.8	3.6
III-1	4	3.0	$\boxed{1.7}$	$\boxed{1.8}$	$\boxed{1.5}$	12.5	19.2	11.4	13.0	17.7	96.2	32.0	3.2
III-2	4	$\boxed{2.2}$	13.0	6.0	15.5	16.1	20.5	15.7	16.1	25.3	81.0	25.8	5.3
III-3	3	$\boxed{0.5}$	4.5	$\boxed{0.7}$	$\boxed{1.7}$	13.9	22.2	10.7	13.8	21.1[b]	110.7	31.4	5.5
IV	15	$\boxed{1.0}$	$\boxed{1.2}$	$\boxed{1.4}$	$\boxed{1.5}$	11.2	15.8	11.2	10.9	15.6	73.3	15.0	2.1

[a] All averages for the II genotypes used as ♀ were weighted according to the number of crosses involved, and consequently the least significant differences were computed according to the average number of crosses within the different male genotypes. On the other hand, the averages for the III and IV genotypes used as ♀ were obtained from the individual plant averages within each class, and consequently each average was based on the same number of variates for computing the least significant differences.

[b] Since this average was based on only two plants, the data for selfing were not included in the analysis for the ♀ genotype III-3.

In order to test the behavior of oppositional alleles in this doubled condition, each of the 29 F_1 plants from the 64 × 64 cross was self-pollinated, crossed reciprocally to both parents, and crossed diallely with all of the sister F_1 plants. The results (table 1) were more complicated than those from the 32 × 32 cross, and they cannot be explained by any of the existing hypotheses. Although the 64-chromosome parents were self-incompatible like their 32-chromosome counterparts, the F_1 generation segregated into three self-incompatible plants and 26 self-compatibles. These 29 plants were self-pollinated by tripping 10 flowers per head in the greenhouse and by manipulating entire heads in both greenhouse and field. The classification into compatible and incompatible types was very distinct by all three methods, and the results from the three methods agreed (table 1). The differences in average seed set between the 26 self-compatible plants were highly significant in each of the three methods, but there were no significant correlations between methods. The value of r was 0.36 when correlating the two greenhouse tests, and it was only 0.10 when correlating the results of manipulation in greenhouse and field.

In order to explain the 3 : 26 segregation, it is proposed that diallelic plants, which produce only one kind of heterozygous pollen, are self-incompatible, whereas triallelic and tetraallelic plants, which produce, respectively, three and six kinds of heterozygous pollen, are self-compatible. It must be assumed: (*a*) When only one type of heterozygous pollen is produced, the presence of one or both of the same factors in a pistil will inhibit that pollen, and (*b*) when more than one type of heterozygous pollen is produced by a plant, none of it is inhibited even if the same factors are present in the pistil. [Assumption (*a*) has been verified by the work of Lewis,[7] but assumption (*b*) is new to this subject; no previous studies have included comparable diallelic, triallelic and tetraallelic genotypes in all possible combinations.] On the other hand, homozygous pollen is assumed to be inhibited whenever the same allele is present in the pistil. Such a proposal is more complex than would be needed to explain only the differences in self-compatibility, but in this form it is also useful in explaining certain cross-incompatibilities. This concept implies that pollen from a tetraploid is not always specifically inhibited by the same genes in the pistil, as is the case with diploids; the ability or inability to grow depends instead on the reaction between the stigma and all the types of pollen placed upon it. Perhaps chromosome doubling has brought into action certain modifying genes that did not express themselves in the diploid, similar to the results obtained by Randolph[8] and Fischer.[9] In crossing the two balanced diallelic parents, which were shown by their 32-chromosome sister cuttings to be different in all four alleles, the F_1 should consist of a 1:8 ratio of self-incompatibles:self-compatibles (i.e., four diallelic:32 triallelic or tetraallelic). The obtained ratio 3:26 fits the expected 1:8

TABLE 2

Theoretical Compatibility Reactions for the Diallel Intercrosses and Reciprocal Backcrosses with an F_1 Population Obtained by Crossing Two Diallelic $4n$ Parents Which Differed in All Four Alleles. A + Sign Indicates That the Pollen Should Function, and 0 That It Should Not. The Reaction with the Homozygous Types Is Shown on the Upper Line in Each Box and the Heterozygous in the Lower. See Text for Hypothesis Explaining Growth of Homozygous and Heterozygous Pollen. All Crosses Included within the Boxes Formed by the Heavy Lines Are Considered Incompatible

	♀ Genotypes with Expected Frequency	Pollen tube reactions of genotypes producing the following gametes:										
		A	B	C	D	E	F	G	H	I	J	K
		1 S_1S_1 1 S_2S_2 4 S_1S_2	1 S_3S_3 1 S_4S_4 4 S_3S_4	1 S_1S_1 1 S_3S_3 4 S_1S_3	1 S_1S_1 1 S_4S_4 4 S_1S_4	1 S_2S_2 1 S_3S_3 4 S_2S_3	1 S_2S_2 1 S_4S_4 4 S_2S_4	1 S_1S_1 1 S_3S_4 2 S_1S_3 2 S_1S_4	1 S_2S_2 1 S_3S_4 2 S_2S_3 2 S_2S_4	1 S_3S_3 1 S_1S_2 2 S_1S_3 2 S_2S_3	1 S_4S_4 1 S_1S_2 2 S_1S_4 2 S_2S_4	1 S_1S_2 1 S_1S_3 1 S_1S_4 1 S_2S_3 1 S_2S_4 1 S_3S_4
A	♀ Parent $S_1S_1S_2S_2$	0 0 0	+ + +	0 + 0	0 + 0	0 + 0	0 + 0	0 + + +	0 + + +	+ + + +	+ + + +	+ + + + + +
B	♂ Parent $S_3S_3S_4S_4$	+ + +	0 0 0	+ 0 0	+ 0 0	+ 0 0	+ 0 0	+ + + +	+ + + +	0 + + +	0 + + +	+ + + + + +
C	1 $S_1S_1S_3S_3$	0 + 0	0 + 0	0 0 0	0 + 0	+ 0 0	+ + +	0 + + +	+ + + +	0 + + +	+ + + +	+ + + + + +

D	1 $S_1S_1S_4S_4$	0 + 0	+ 0 0	0 + 0	$\boxed{\begin{smallmatrix}0 & 0\\ 0\end{smallmatrix}}$	+ + +	+ 0 0	0 + + +	+ + + +	+ + + +	0 + + +	+ + + + + +
E	1 $S_2S_2S_3S_3$	+ 0 0	0 + 0	+ 0 0	+ + +	$\boxed{\begin{smallmatrix}0 & 0\\ 0\end{smallmatrix}}$	0 + 0	+ + + +	0 + + +	0 + + +	+ + + +	+ + + + + +
F	1 $S_2S_2S_4S_4$	+ 0 0	+ 0 0	+ + +	+ 0 0	0 + 0	$\boxed{\begin{smallmatrix}0 & 0\\ 0\end{smallmatrix}}$	+ + + +	0 + + +	+ + + +	0 + + +	+ + + + + +
G	4 $S_1S_1S_3S_4$	0 + 0	$\boxed{\begin{smallmatrix}0 & 0\\ 0\end{smallmatrix}}$	$\boxed{\begin{smallmatrix}0 & 0\\ 0\end{smallmatrix}}$	$\boxed{\begin{smallmatrix}0 & 0\\ 0\end{smallmatrix}}$	+ 0 0	+ 0 0	0 + + +	+ + + +	0 + + +	0 + + +	+ + + + + +
H	4 $S_2S_2S_3S_4$	+ 0 0	$\boxed{\begin{smallmatrix}0 & 0\\ 0\end{smallmatrix}}$	+ 0 0	+ 0 0	$\boxed{\begin{smallmatrix}0 & 0\\ 0\end{smallmatrix}}$	$\boxed{\begin{smallmatrix}0 & 0\\ 0\end{smallmatrix}}$	+ + + +	0 + + +	0 + + +	0 + + +	+ + + + + +
I	4 $S_1S_2S_3S_3$	$\boxed{\begin{smallmatrix}0 & 0\\ 0\end{smallmatrix}}$	0 + 0	$\boxed{\begin{smallmatrix}0 & 0\\ 0\end{smallmatrix}}$	0 + 0	$\boxed{\begin{smallmatrix}0 & 0\\ 0\end{smallmatrix}}$	0 + 0	0 + + +	0 + + +	0 + + +	+ + + +	+ + + + + +
J	4 $S_1S_2S_4S_4$	$\boxed{\begin{smallmatrix}0 & 0\\ 0\end{smallmatrix}}$	+ 0 0	0 + 0	$\boxed{\begin{smallmatrix}0 & 0\\ 0\end{smallmatrix}}$	0 + 0	$\boxed{\begin{smallmatrix}0 & 0\\ 0\end{smallmatrix}}$	0 + + +	0 + + +	+ + + +	0 + + +	+ + + + + +
K	16 $S_1S_2S_3S_4$	$\boxed{\begin{smallmatrix}0 & 0\\ 0\end{smallmatrix}}$	$\boxed{\begin{smallmatrix}0 & 0\\ 0\end{smallmatrix}}$	$\boxed{\begin{smallmatrix}0 & 0\\ 0\end{smallmatrix}}$	$\boxed{\begin{smallmatrix}0 & 0\\ 0\end{smallmatrix}}$	$\boxed{\begin{smallmatrix}0 & 0\\ 0\end{smallmatrix}}$	$\boxed{\begin{smallmatrix}0 & 0\\ 0\end{smallmatrix}}$	0 + + +	0 + + +	0 + + +	0 + + +	+ + + + + +

very closely with a non-significant corrected $\chi^2 = 0.031$. Similarly, it fits the expected 9:40 ratio, based on random chromatid assortment, with a non-significant corrected $\chi^2 = 0.767$.

When the three self-incompatible F_1 plants were diallely inter-crossed and reciprocally back-crossed to both parents, their differential reaction allowed them to be divided into two groups (table 1). One of the three plants (Group II-1) was compatible as female with its female parent but incompatible with its male, while the other two (Group II-2) were incompatible as females with their female parent, compatible with their male, and cross-incompatible. Although the crosses II-1 × II-♀ P, II-2 × II-♂ P and II-2 × II-1 are considered compatible, it should be pointed out that in every case the average seed set was significantly lower than that obtained in the other compatible crosses of these same three plants as females with their 26 self-compatible sisters.

All crosses involving the self-compatible plants as males were compatible, but several differential reactions occurred among the crosses self-compatible ♀ × self-incompatible ♂ (table 1). Four self-compatible plants (Group III-1) set significantly more seed when crossed in this direction with their female parent than they did with the other self-incompatible plants. This average of 3.0 with the female parent, however, showed a highly significant difference in comparison with the compatible crosses between the same four plants as females and their self-compatible sisters as males. Four other self-compatible plants (Group III-2) proved incompatible with their female parent when the latter was used as male but compatible with the other three self-incompatible groups used as males. In this case, the cross III-2 × II-1 has been classed as compatible, but its average of 6.0 was significantly different from all other compatible crosses using Group III-2 as female. A third group of three self-compatible plants (Group III-3) set an average of 4.5 seeds per head when crossed as females with their male parent, and the differences between this and the crosses with the other self-incompatible groups as males were highly significant. Likewise, the differences between 4.5 and the other compatible crosses using Group III-3 as female were highly significant. The remaining 15 self-compatible plants (Group IV) proved incompatible as females with both parents and with both self-incompatible F_1 groups. The differences between the four types of incompatible crosses involving Group IV as female were not significant.

If pollen-tube growth in crosses is conditioned in the same manner as was proposed for selfings, the theoretical pattern of incompatibilities (table 2) is similar to that which was obtained (table 1). The segregation of 1:2:4:4:3:15 in the F_1 population gives a very satisfactory fit to expected, with a non-significant $\chi^2 = 4.904$ (7 D.F.). When the same ratio was fitted to the expected based on independent chromatid assortment, a non-significant $\chi^2 = 7.805$ was obtained. Probably because the F_1 population was small, only two self-incompatible groups were found. In order to fit these results to the pattern of cross-incompatibilities obtained with the self-compatible plants as females, it is apparent that the two self-incompatible groups which were obtained were either C and E or D and F (table 2). The data do not fit in all respects, however, and certain additional assumptions must be made. One discrepancy from the expected results was that those self-compatible plants (Group III-1) which were compatible as females with their female parent, but incompatible with their male parent should have been incompatible as females with one or the other of the self-incompatible F_1 groups but not with both. Each of the four plants in this group gave some indication of setting more seed with one self-incompatible F_1 group than with the other, but the differences were neither large enough nor consistent enough to be significant. The second and more serious deviation from expected was the incompatibility obtained between each of the self-incompatible F_1 plants as female and one of their parents. According to the hypothesis, these crosses should have been compatible, and no adequate explanation of this irregular behavior is available at present.

Two additional lines of evidence have been obtained which support the concept of pollen-tube growth sometimes depending on the interaction between the stigma and all pollen grains placed upon it rather than on the specific relation between the alleles borne by the pollen and the stigma. In the first place, significant differences in average seed set were obtained from the different male groups. No such effect was observed in the 32 × 32 chromosome test, nor has it ever been observed before in white clover in the thousands of crosses which have been made with normal 32-chromosome plants. Several cases of morphological or cytological abnormalities causing defective pollen and consequent male sterility have been observed in 32-chromosome clover, but no differences due to genetic causes, where there was no influence on pollen fertility, have been demonstrated. Among the 64 × 64 matings, some of the most outstanding cases have already been pointed out. Further examples may be seen (table 1) by comparing the differences between group averages with the least significant differences. Similar differences were found when the results were grouped in other ways. For example the average of all crosses using the triallelic plants as males was 15.4, while the average for the tetraallelic as males was 13.4, and the difference between these averages was highly significant. Similarly, the triallelic plants set significantly more selfed seed (av. = 21.4) than the tetraallelic plants (av. = 15.6). One possible explanation of these relationships might be that more complex interactions and somewhat lowered seed set would result from six types of heterozygous pollen on the stigma than from three; further tests are now being made. With these data it was possible to demonstrate in addition highly significant differences between the plants within each of the groups in regard to their differential male influence. It was also noticed that with most of the self-compatible plants, the seed set from selfing was significantly greater than from crossing. Most of the selfs were made without emasculation, whereas the crosses were made following emasculation, but differential treatment of this sort has never given different results with normal 32-chromosome plants. It is necessary to postulate modifying factors to explain such behavior, and there may be a better chance for their expression in tetraploid plants.

The other additional line of evidence supporting the interaction hypothesis was derived from the significant end-season increases in seed set obtained in the green house. This behavior likewise has not been observed in 32-chromosome white clover. When it was noticed that the reflection of pedicels was not so definite a reaction with the 64-chromosome plants as it was with the 32, 103 crosses were made in duplicate. The first crosses of each pair averaged 13.8 seeds per head, while the second averaged 17.2. The average interval between them was 10.9 days. The differences between repeats were analyzed for each of the eight types of compatible matings between diallelic, triallelic and tetraallelic plants, but significant variance due to repeats was found only for the crosses IV × III and IV × IV. These combinations would involve the greatest number of interactions between pollen and pistil, and therefore might show the greatest influence from environmental factors such as end-of-season.

Summary.—The 64-chromosome stolons of white clover, induced by colchicine treatment of seedlings, were found to be self-incompatible like the 32-chromosome plants from which they arose. In the F_1 population from two of the 32-chromosome plants, four intra-sterile, inter-fertile groups were found, and all F_1 plants were reciprocally compatible with both parents, indicating that the parents differed in all four oppositional alleles. In the F_1 from the two corresponding 64-chromosome stolons, three plants were self-incompatible, and the remaining 26 were self-compatible. All intercrosses involving the self-compatible plants as males were compatible, but both compatible and incompatible matings were found when the self-incompatible parents and F_1 plants were used as males. On the basis of these differential reactions, several genotypes were postulated. It is suggested that growth of pollen bearing two different alleles depends sometimes on the interaction between the stigma and all pollen placed upon it rather than on a specific oppositional effect.

---

[1] Contribution No. 58, of the U. S. Regional Pasture Research Laboratory, Division of Forage Crops and Diseases, Bureau of Plant Industry, Soils, and Agricultural Engineering, Agricultural Research Administration, U. S. Department of Agriculture, State College, Pennsylvania, in coöperation with the northeastern states.

[2] East, E. M., and Mangelsdorf, A. J., *Proc. Nat. Acad. Sci.*, **11**, 166–171 (1925).

[3] Stout, A. B., and Chandler, C., *Science*, **94**, 118 (1941).

[4] Stout, A. B., and Chandler, C., *Ibid.*, **96**, 257–258 (1942).

[5] Crane, M. B., and Lewis, D., *Jour. Genetics*, **43**, 31–43 (1942).

[6] Howard, H. W., *Nature*, **149**, 302 (1942).

[7] Lewis, D., *Jour. Genetics*, **45**, 171–185 (1943).

[8] Randolph, L. F., *Amer. Nat.*, **75**, 347–363 (1941).

[9] Fischer, H. E., *Genetics*, **26**, 151 (1941).

[10] Atwood, S. S., *Jour. Amer. Soc. Agron.* (in press).

[11] Atwood, S. S., *Ibid.*, **33**, 538–545 (1941).

LINKAGE STUDIES OF THE RAT (RATTUS NORVEGICUS) VI.

BY W. E. CASTLE AND HELEN DEAN KING

UNIVERSITY OF CALIFORNIA AND WISTAR INSTITUTE

Communicated February 28, 1944

1. *A Fourth Linkage Group.*—In a canvass of the linkage relations of genes of the rat, it has been found that *kinky* and *stub* are borne in a common chromosome. The evidence is as follows. Kinky (*k*) is a recessive mutant gene discovered by Feldman.[1] It makes the hair short and curly, being similar in its phenotypic effects to the two dominant mutations curly and curly$_2$. Stub (*st*) is a mutation likewise recessive and frequently lethal at an early age, discovered by King.[2] It makes the tail short and "stubby" and frequently results in abnormality of the kidneys, and other organs in the pelvic region, the hind legs being sometimes partially fused. Microphthalmia also is often observed, particularly of the left side. Since stub animals, even when they live to normal breeding age, are usually sterile, only animals heterozygous for the stub gene are available for genetic study. When they are crossed with animals homozygous for kinky, half the resulting progeny will be heterozygous for both genes, though they show neither. Such carriers can be identified by suitable test matings. They should theoretically produce gametes of four classes, *k*, *st*, *k st* and *O*. The first two would be non-crossover (repulsion) gametes, the last two would be crossover (recombination) classes. To ascertain the relative frequency of their occurrence is a task best undertaken by an outcross of an F_1 animal to one which carries neither gene. The resulting progeny may at maturity be tested individually by a backcross to an F_1 animal known to be a carrier of both *k* and *st*. Usually it is most convenient to employ out-crossed females in back-crossing to tested F_1 males. A total of 138 out-crossed females have thus been tested. They fall into the four expected classes with frequencies as indicated, $\frac{k}{66}\,\frac{st}{25}\,\frac{k\ st}{11}\,\frac{O}{36}$; crossovers $11 + 36 = 47$; non-crossovers $66 + 25 = 91$. Total number 138. Deviation from equality

of the two groups 22, P.E. 3.96. Dev./P.E. = 5.5, a significant and clear demonstration of repulsion between the two genes. Crossover percentage 34.1 ± 2.6.

It will be observed that of the 138 tested animals 77 are recorded as carriers of the *k* gene. This is a little more than the expected half. But only 36 tested animals are recorded as carriers of *st*, which indicates that many such failed of identification in the test matings through the production of too small a number of young (six or more being rated a satisfactory test) or because the stub young died at or before birth and were eaten before the test litters were inspected.

Consequently it would seem desirable to correct the deficiency of stub carriers by increasing the number to half the total population, and decreasing correspondingly the classes recorded as lacking *st*. The deficiency of stub carriers is 33, requiring an increase of 16 each in classes *k st* and *st*, and a corresponding decrease in classes *O* and *k*. Making these changes the classes would become $\frac{k}{50}\,\frac{st}{41}\,\frac{k\ st}{27}\,\frac{O}{20}$, a more symmetrical distribution, though with an unchanged crossover percentage, 34.1.

Before indications of linkage between stub and kinky had been obtained, tests had been made with entirely negative results for linkage of stub with wobbly (*wo*), with blue dilution (*d*), and with Curly$_2$ (Cu_2).

2. *A Further Study of the Second Linkage Group.*—In earlier studies[8] it has been shown that the dominant gene curly is loosely linked with the recessive and lethal gene anemia (*an*) and the recessive gene incisorless (*in* complete absence of incisor teeth).

To get a better knowledge of the order of the genes and their map distances in the curly region, a three-point cross has been made between curly animals (carrying anemia) and incisorless animals. The resulting F_1 animals were in genetic formula $\frac{Cu\ an\ +}{+\ +\ in}$. They were out-crossed to animals free from all three mutant genes, and the out-crossed progeny were then back-crossed to F_1 to ascertain how many and which of the three genes each of them carried. A total of 175 animals have been thus tested, and they fall into 8 classes. Non-crossover classes are represented by the genotypes *Cu an*, 65 individuals, and *in*, 76 individuals. The six crossover genotypes and their frequencies are as follows:

Cu in	7	*An in*	3
an	3	*Cu an in*	5
Cu	5	*O*	11, total 34.

It was shown in an earlier publication that *an* probably lies between *Cu* and *b* (brown) though much closer to the former. If the order of the three genes is as assumed *Cu an b*, the position of *in* with reference to *Cu an* may be either (1) *Cu an in* or (2) *in Cu an* or (3) *Cu in an*. Let us consider the evidence for each of these possible arrangements.

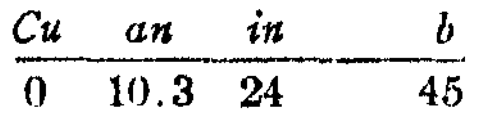
Whatever the order of the three genes, the total occurrences of crossing over (single or double) in the 175 cases studied will be the same, 34. But the number of double crossovers will be different under each assumed arrangement. Thus it will be 8 under order (1), 10 under order (2) and 16 under order (3).

ORDER	SINGLE	DOUBLE	% DOUBLE
(1) *Cu an in*	10 + 16 = 26	8	23.5
(2) *in Cu an*	16 + 8 = 24	10	29.4
(3) *Cu in an*	10 + 8 = 18	16	47.0

That order is most probable under which double crossovers are least absolutely and relatively. This points to order (1) as most probable and order (3) as least probable. Order (3) is further improbable because it would make the crossover percentage between *Cu* and *an* 28.5, whereas an earlier study in which *in* was not involved (and in which consequently the apparent crossing over between *Cu* and *an* would be less) showed only 2.1 ± 0.5%. Under either order (1) or order (2) the value for *Cu an* would in this investigation be 10.3%, a much smaller discrepancy.

Assuming that order (1) is correct, the total amount of crossing over in the region *Cu an* will be 10 single + 8 double = 18 = 10.3 ± 1.6%. In the region *an in* there will be 16 single + 8 double = 24 = 13.7 ± 1.7%. Consequently the map for the second chromosome may be drawn thus:

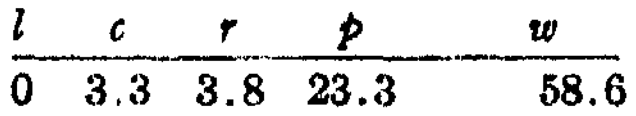
II.



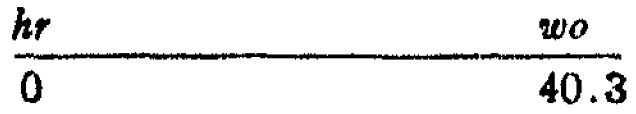
A test of the validity of this arrangement by a study of the linkage relation of *in* to *b* is highly desirable.

For comparison we may add here the other linkage maps of the rat so far as at present known.

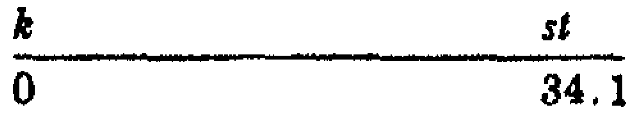
I.



III.



IV.



[1] *Jour. Hered.*, **26**, 162 (1935).

[2] *The American Anatomical Memoirs*, No. 17, p. 59 (1939); *Anat. Rec.* **81**, p. 283 (1941).

[3] *Proc. Nat. Acad. Sci.*, **21**, 390 (1935); Ibid., **26**, 578 (1940). *Carnegie Institution Year Book*, No. 41, p. 225 (1941–1942).

EQUILIBRIUM IN GENIC MATERIALS

By Donald F. Jones

Connecticut Agricultural Experiment Station

Communicated February 25, 1944

Inbred strains of maize, continuously inbred for many generations, are favorable material to show small heritable variations. Changes in cob and silk color, plant and ear size, and in the amount of chlorophyll have been reported (Jones, 1924, 1939; Singleton, 1943). These alterations either have no effect upon vigor and productiveness or clearly reduce the plants in some degree. In addition to these, other heritable changes have been found recently in various inbred material. All of these lessen the ability of the plant to grow and to reproduce itself in some measure. In their total effect upon the inbred plants, they appear to be degenerative changes.

Unexpectedly, it has been found that these apparently degenerate lines, when crossed back with their parental lines, show heterosis. Presumably, they differ by only a single gene. This point remains to be tested. Furthermore, one of these degenerate lines has been shown to give better results, than the parental normal line, in outcrosses. The changes so far noted affect both physiological and morphological characters and are listed as follows:

INBRED LINE	CHARACTERISTIC FEATURE	NATURE OF CHANGE
C14[*]	Late flowering	Tassels and silks appear later
C14	Blotched leaf	Chlorophyll reduced
C30	Reduced plant	Smaller in all parts, blooms earlier
C237	Reduced ear	Ears smaller, stalks taller
Illinois Hy	Dwarf plant	Internodes reduced in length
C.I. 4–8	Narrow leaf	Leaf width and stalk diameter smaller, blooms earlier
C.I. Kr (187-2)	Crooked stalk	Internodes shortened and bent
C.I. Kr (187-2)	Pale top	Chlorophyll reduced

[*] The letter C designates inbreds produced at the Connecticut Agricultural Experiment Station, New Haven, Connecticut. Illinois refers to the Agricultural Experiment Station at Urbana, Illinois, and C.I., the Office of Cereal Investigations, U. S. Department of Agriculture, Washington, D. C.

The reduced plant, C30, has been described by Singleton (1943, 1944). It is a change in the Purdue 39 inbred sweet corn originally found in Idaho. It is used in the production of Connecticut hybrid sweet corn. This single gene mutation gives evidence for increased growth in crosses with unrelated lines as well as a hybrid vigor effect when crossed with the parental line.

Increased growth in crosses of related lines was noted during the 1943 growing season in the narrow leaf line crossed back to the original plants from which it was derived. The narrow leaf condition was found two years previously in one progeny that had been grown from a single self-fertilized ear. When grown the following year in adjoining rows it was seen to be clearly distinct in width of leaf. It was crossed back to the original line to see if it would segregate as a single factor difference. When the original line, the deviating line and their first generation hybrid were all grown in adjoining rows in 1943, differences were clearly noticed.

The originally 4-8 line and the same line from a new source were alike, showing that no visible change had occurred in the parental line. The deviating narrow leaf line was smaller in leaf width and stalk diameter, the same or slightly taller in height of stalk, the same in time of shedding pollen but earlier by seven days in time of silking. Thus it was not a degeneration in all respects as earliness of flowering is a good measure of growth efficiency in maize. Any environmental or inherited condition that favors growth is reflected in a speeding up of the time of flowering.

The first generation hybrid plants resulting from the cross of the deviating narrow leaf line and the original line from which it came was equal to or better than the better parent in all characters. The combined result was a noticeably greener, and more robust plant that was early maturing and more productive of grain. Not enough plants were grown to measure these small differences accurately. The hybrid did not surpass either parent in any single character by a statistically significant margin but the total result was impressive, especially in yield of grain.

Illinois Hy is an inbred that is widely used in the production of hybrid field corn. A dwarf form of this inbred was compared with the original line for two years. It has proved to be typical Hy in all details of color and structure except that the internodes are shorter, and the leaves smaller. The plant is brachytic in type, averaging about half the height of the original line. The dwarf plants flower at the same time as their normal parent and produce smaller ears with smaller kernels and are somewhat less productive in total amount of grain. While the segregation of this character has not been tested, dwarfs of this type have appeared in other lines and are usually due to a single mendelizing recessive gene.

The first generation of the cross of dwarf and normal Hy was grown alongside the normal plants. The crossed plants averaged the same in height but shed pollen two days earlier and silked four days earlier. The plants were noticeably greener and more thrifty in appearance during the growing season and at maturity the ears were larger with brighter and better filled kernels. As the hybrid vigor effect was not anticipated only a few plants were grown and most of the ears were used for hand pollination so that no yield figures are available. The earlier flowering and maturing is a clear indication of heightened growth efficiency.

An inbred line of Leaming, C14, has been continuously self-fertilized for 21 generations. After the first ten generations it was reduced to a high degree of uniformity and constancy. In the 17th generation a conspicuously tall darker green plant, flowering early, was noted in one progeny. This single plant had the appearance of a typical C14 slightly enlarged and invigorated as if it was growing in a particularly fertile spot. It was not large enough to be an outcross such as appear occasionally as a result of contamination with pollen from unrelated plants. This deviating plant was self-fertilized and the resulting ear was larger and better matured than the other self-fertilized ears in the same progeny. The kernels were marked with a faint pericarp stripe not characteristic of any other varieties that had been grown that year or before.

The progeny grown from this ear was noticeably taller, greener and more productive than the other related lines. It was slightly more variable but showed no clear segregation. Several plants were self-fertilized and grown for two successive generations. Some of these now differ from the original line in glume and internode color and in the amount of chlorophyll blotching shown by the leaves late in the season. All progenies have been reduced to the same level of productiveness as the original line or below.

The segregation of color markings indicates that the original plant could have resulted from an outcross with a related back-crossed line that was growing nearby the year before the deviating plant was discovered. It was so considered until the behavior of the other deviating lines was noted. The possibility that this is a mutation originating in the heterozygous condition must now be considered.

One line from this source more heavily blotched than the original is now noticeably less vigorous and productive. Whatever its origin it differs from the original C14 line by only a small germinal alteration. When crossed with the normal line the first generation hybrid plants were taller, greener and more productive. Here again the differences are small and not statistically significant for any one measurable character but all are in the same direction of increased growth efficiency.

The reduced ear size character has been described previously (Jones, 1939) as a change in two directions. The number of kernels on the ear is reduced but the height of plant is increased. Crossed with the original this deviating line showed no increasing effect in either height of stalk or production of grain. The first generation plants were intermediate in both respects.

A single factor mutation in tobacco (Jones, 1921) gives results in the heterozygous condition that are superior in some respects to either parent. In this case the change is from the determinate to the indeterminate habit of growth. The heterozygous plants average several more leaves and more flowers per plant than the determinate parent and produce a larger amount of total plant material than either parent in the same period of growth.

Several other deviating lines in maize have been found but have not yet been compared in crosses with their parental lines. Late flowering plants appeared in another progeny of C14 also in the 17th generation. The seedlings germinate more slowly and the plants are shorter than their normal sibs throughout the growing period. At the end of the season, when the normal plants cease to increase in height, the late flowering plants continue growth for a longer time until they are either as tall or almost as tall as normal when elongation ceases.

In 1942 the normal line averaged 84.0 inches in height to the tip of the tassel and the late flowering line 78.9. In 1943 the heights are 78.2 and 75.9, respectively. In time of flowering in 1943 the two lines compare as follows:

DAYS FROM PLANTING TO	FIRST TASSEL	LAST TASSEL	FIRST SILK	LAST SILK
Late flowering line	63	67	66	78
Normal flowering line	59	65	61	66
Differences	4	2	5	12

The number of days is counted from date of planting to pollen shedding and first appearance of silks. The results are based on an average of four replicated plots in each case. The two lines differ in no morphological character and at the end of the season appear to be alike in every respect.

In 1942 several progenies of C.I. Kr (187-2) were grown from single self-fertilized ears all coming from a single line that had been inbred for many years and grown in Connecticut four generations with continuous self-fertilization. No unusual variations had been noted in previous years. In 1942 several progenies were segregating for an obscure chlorophyll variation. A few plants were noted in the early stages of growth as being shorter and lighter in color at the base of the upper leaves. These plants were marked and self-fertilized. At time of flowering the differences had largely disappeared, the plants were entirely normal in appearance but were usually a few days later in silking. At maturity the plants were so nearly the same as the normal plants in height and productiveness that they could not be noted as deviations.

When grown the following year two progenies were homozygous for this pale top condition, four were segregating and six were free of it. In the homozygous progenies the plants were now noticeably reduced in size throughout the season. The stalks were more slender and shorter at maturity. One homozygous pale top progeny measured 72.8 inches in height compared to six normal progenies that averaged 76.8. The plants flowered later and the ears were smaller and less well matured. Chlorophyll color was a little lighter throughout the season although the plants grew out of the early pale top condition and were not clearly differentiated in this respect after flowering. Only when they were growing together in an entire row were they distinctly different. Nevertheless, the deviating plants appeared to be more differentiated from normal the second year than when they first appeared as segregates from normal. The condition had become progressively worse.

In one of the progenies, segregating for pale top in 1942, one normal green plant was noted as distinctly shorter at maturity. Upon closer examination it was seen that two internodes at the base of the plant were much shortened and bent. The progeny grown from the self-fertilized seed of this plant were all crooked in the same way as the parent, varying in the number of internodes affected from one to five. In some plants alternate internodes are shortened and bent, the intervening internodes being normal.

Out of thirteen progenies grown, six are normal, four are segregating normal and crooked stalk, and three are all crooked. In two of the latter the parental plant was not noted as being abnormal in stalk formation. All of the progenies showing crooked stalks are also either homozygous or heterozygous for pale top and all of the progenies showing pale top are either homozygous or heterozygous for crooked stalk. Apparently the two deviating types originated at the same time. They both behave as mendelian recessives. One progeny is also segregating for flat cobs. Further tests are needed to show whether or not these three conditions are associated in any way and the meaning of the simultaneous appearance of the three degenerate types.

Neither pale top nor crooked stalk have been tested in crosses with normal related or unrelated lines. But there is an indication that the normal plants heterozygous for one or both conditions are taller than homozygous normals. Two progenies segregating for pale top and crooked stalks averaged 80.1 inches in height compared to the six normal progenies which averaged 76.8 in adjacent rows. The segregating progenies had several deviating plants that were clearly shorter than normal and these were included in the measurements. The heterozygous normal plants must be taller than the homozygous normal to more than make up for this difference.

In all cases so far observed the deviating lines originate in heterozygous normal plants that are selected either wittingly or unwittingly for propagation. The larger growth of the heterozygous plants increases the chance that they will be used as the progenitors for the succeeding generations. The maintenance of inbred strains to be used in the production of hybrid corn is a more difficult problem than was at first realized. Even if the reduced lines do give better results in crosses the propagation of the inbred line may become prohibitively costly. This apparently has happened in several widely used inbreds which are now so unproductive as to be difficult to produce in those regions where the hybrids from them are still among the most productive.

The increased growth shown by some and possibly all of these related crosses results in each case from a single allelic difference. This at first sight appears to be a stimulus of heterozygosis as originally postulated by East (1909), East and Hayes (1912) and Shull (1914). This conception was extended by Rasmusson (1934) and by East (1936) as an interaction between alleles. East dismissed these defective genes as having no importance in heterosis. It is now apparent that no clear distinction can be made between normal and degenerate characters since they may have multiple effects, some favorable, some unfavorable. Dunn and Caspari (1942) describe numerous changes in a small section of a mouse chromosome having diverse effects, some lethal, others having degenerate effects. Hybrid combinations of some of these multiple changes promote normal growth.

Deviating lines show a loss in some respects and a gain in others. The earlier flowering of the narrow leaf and reduced plant lines is an indication of increased growth efficiency. There seems to be an equilibrium in genic substances such that a change in one direction may be counter-balanced by a change in another. Whatever the deviating line has lost the original line supplies in their crossbred offspring. What the deviating line has acquired may be added to the sum total of genic action. Along with this there is the accumulation of non-allelic, favorable growth factors that is known to occur.

This investigation has been aided by a grant from the Rockefeller Foundation.

Dunn, L. C., and Caspari, E., these PROCEEDINGS, **28**, 205–210 (1942).
East, E. M., *Amer. Nat.*, **43**, 173–181 (1909).
East, E. M., *Genetics*, **21**, 375–397 (1936).
East, E. M., and Hayes, H. K., *U. S. Dept. of Agr., Bur. Plant Ind. Bull.*, **243** (1912).
Jones, D. F., *Genetics*, **6**, 433–444 (1921).
Jones, D. F., *Ibid.*, **9**, 405–418 (1924).
Jones, D. F., *Ibid.*, **24**, 462–473 (1939).
Rasmusson, J., *Hereditas*, **18**, 245–261 (1934).
Shull, G. H., *Zeit. ind. Abst. vererb.*, **12**, 97–149 (1914).
Singleton, W. R., *Genetics*, **28**, 89 (1943).
Singleton, W. R., *Records of Genetics Society of America for 1943*, (1944).

ON THE RÔLE OF CARBON DIOXIDE IN THE METABOLISM OF CLOSTRIDIUM THERMOACETICUM

By H. A. Barker

Division of Plant Nutrition, University of California

Communicated March 2, 1944

Clostridium thermoaceticum ferments glucose anaerobically with the formation of somewhat more than 2.5 moles of acetic acid per mole of glucose decomposed.[3] Besides acetic acid, only cell material and a trace of carbon dioxide are produced. The remarkable features of this "acetic acid fermentation" are (1) the formation of more than two moles of acetic acid per mole of glucose and (2) the absence of carbon dioxide or other C_1 compound which would be expected to arise by a C_1-C_2 split of a C_3 intermediate.

In explanation of the high yield of acetic acid and the absence of carbon dioxide, two hypotheses have been suggested.[3] One is the primary cleavage of glucose into three C_2 fragments. The other is the production and later reabsorption of carbon dioxide. The latter hypothesis seems much more probable in view of the fact that two other anaerobic spore-formers are already known to convert carbon dioxide into acetic acid. The rôle of carbon dioxide in the metabolism of these bacteria is that of an oxidizing agent. *Clostridium aceticum*[4, 5] uses it for the oxidation of molecular hydrogen according to the equation

$$4H_2 + 2CO_2 = CH_3COOH + 2H_2O$$

while *Clostridium acidi-urici*[1, 2] uses it for the oxidation of uric acid and other purines. The acetic acid fermentation of *C. thermoaceticum* may similarly be regarded as an anaerobic oxidation of glucose by means of carbon dioxide as expressed by the equations

$$\begin{array}{lll} \text{Oxidation:} & C_6H_{12}O_6 + 2H_2O & = 2CH_3COOH + 2CO_2 + 8H \\ \text{Reduction:} & \underline{8H + 2CO_2} & \underline{= CH_3COOH + 2H_2O} \\ \text{Net reaction:} & C_6H_{12}O_6 & = 3CH_3COOH \end{array}$$

There are two direct methods by which the above hypothesis regarding the rôle of carbon dioxide in the acetic acid fermentation might be tested. One involves the use of carbon dioxide labeled with a carbon isotope. This method cannot at present be applied because of wartime restrictions and the recent untimely death of Dr. S. Ruben who was planning to cooperate in these experiments. The second method consists of studying the action of the bacteria on a substrate more reduced than glucose, such as glycerol or a sugar alcohol. Since glucose is fermented without carbon dioxide evolution, it follows that the decomposition of a more reduced substrate must involve carbon dioxide uptake if carbon dioxide is acting as an oxidizing agent. Unfortunately this method also cannot be applied to the problem at hand because the strain of *C. thermoaceticum* at present available does not attack any of the polyalcohols.

Since no direct method of detecting carbon dioxide utilization is available, an indirect approach based upon the following line of reasoning has been used. If glucose decomposition involves a splitting of the molecule into three C_2 pieces (the first hypothesis mentioned above) the decomposition of a pentose should yield two C_2 and one C_1 fragments. If, on the contrary, acetic acid is synthesized entirely or in part from carbon dioxide, it should be the only product of a pentose as well as a hexose fermentation. The experimental results (table 1) show that except for traces of carbon dioxide, acetic acid is the only product formed from *l*-xylose. Therefore a utilization of carbon dioxide is indicated.

TABLE 1

FERMENTATIONS OF *l*-XYLOSE AND PYRUVATE BY *Clostridium thermoaceticum*

(mM per 100 mM substrate fermented)

	SUBSTRATES	
PRODUCTS	*l*-XYLOSE	PYRUVATE
Carbon dioxide	2.3	51.3
Acetic acid	223	109
Carbon recovery (%)	90	89
Redox index	Indet.	1.03

Although the results obtained with xylose definitely favor the hypothesis of carbon dioxide reduction, other explanations are, of course, possible. Instead of being directly split into two C_2 and one C_1 fragments, xylose might be broken into a C_2 and a C_3 piece, two of the latter then being converted into a C_6 compound which could be subsequently split into three C_2 fragments. In this way the facts could be explained without assuming carbon dioxide to be either formed or utilized. In view of this theoretical possibility it seemed desirable to find out if *C. thermoaceticum* is able to form carbon dioxide from an appropriate substrate and to see whether indications of acetic acid synthesis could be simultaneously obtained. The decomposition of pyruvate was therefore studied.

Since pyruvate is more oxidized than a carbohydrate, its decomposition must yield carbon dioxide or some other compound more oxidized than acetic acid. Carbon dioxide and acetic acid are in fact the only products. Now the decomposition of pyruvate would be expected, in the absence of carbon dioxide reduction and the synthesis of C_6 compounds, to yield not more than 100 mM of acetic acid per 100 mM of pyruvate decomposed and not less than 100 mM of carbon dioxide per 100 mM of acetic acid formed. Actually the experimental results (table 1) agree rather well with the equation

$$4CH_3COCOOH + 2H_2O = 5CH_3COOH + 2CO_2.$$

The yield of acetic acid is significantly above 100 mM while the yield of carbon dioxide is only about 50 mM per 100 mM of pyruvate fermented. These results are again entirely consistent with the view that carbon dioxide is converted into acetic acid. But proof of this mechanism must await the application of one of the two direct methods mentioned above.

[1] Barker, H. A., Ruben, S., and Beck, J. V., *Proc. Nat. Acad. Sci.*, **26**, 477–482 (1940).

[2] Barker, H. A., and Beck, J. V., *Jour. Biol. Chem.*, **141**, 3–27 (1941).

[3] Fontaine, F. E., Peterson, W. H., McCoy, E., Johnson, M. J., and Ritter, G. J., *Jour. Bact.*, **43**, 701–715 (1942).

[4] Wieringa, K. T., *Leeuwenhoek*, **3**, 1–11 (1936).

[5] Wieringa, K. T., *Ibid.*, **6**, 251–262 (1939–1940).

THE MATRICES OF FACTOR ANALYSIS

By A. A. Albert

Department of Mathematics, University of Chicago

Communicated February 25, 1944

In factor analysis a study is made of n-rowed symmetric matrices M_0 whose *diagonal elements* are *all zero* and whose nondiagonal elements are correlation coefficients, that is, real numbers between zero and unity. A general problem[1] is then to determine a diagonal matrix D such that

$$M = M_0 + D$$

has minimum rank ρ. The diagonal elements of D are those of the factor matrix M and are called the *communalities*. As soon as D is evaluated a standard computational technique may be employed to write M as a product

$$M = FF',$$

where F has n rows and ρ columns,[2] F' is its transpose.

Professor L. L. Thurstone of the Department of Psychology of the University of Chicago has called my attention to this problem with particular emphasis on a special case. This case is of the greatest importance because the matrices which arise from actual psychological tests appear to be of the type assumed for this case.

Our special problem is the study of the case where ρ is equal to what we shall call the *ideal rank* r. We define r to be the largest order of a nonvanishing minor of M_0 obtained by a selection of r of its rows and r *different* columns. No element of such a minor is a diagonal element of M_0 and hence all minors of the kind described are minors of M as well as of M_0. It follows that

$$\rho \geqq r,$$

and that our special problem is a study of the case where ρ attains its minimum value r. It should be clear that

$$m = n - 2r \geqq 0.$$

The procedure we shall give then begins with a computation of the ideal rank and will yield not only r but also the corresponding r row and column labels. Then we may alter our notations so that the columns are the *first* r columns and the rows the *second* r rows. This yields a partitioning of M_0 so as to yield the notation

$$M_0 = \begin{pmatrix} A_0 & B' & C' \\ B & C_0 & H' \\ G & H & K_0 \end{pmatrix} \qquad (1)$$

for M_0. Here A_0, B and C_0 are r-rowed square matrices and the determinant $|B| \neq 0$. The matrices A_0, C_0, K_0 are symmetric matrices with zero diagonal elements, and K_0, G and H have m rows, G and H have r columns. Let us assume that the computation of r has been carried out so as to yield the matrix

$$E = B^{-1}.$$

We now write

$$M = M_0 + D = \begin{pmatrix} A & B' & G' \\ B & C & H' \\ G & H & K \end{pmatrix},$$
$$A = A_0 + X, \qquad B = B_0 + Y, \qquad K = K_0 + Z \qquad (2)$$

for diagonal matrices

$$X = \text{diag.}\{x_1, \ldots, x_r\}, \qquad Y = \text{diag.}\{y_1, \ldots, y_r\}, \qquad Z = \text{diag.}\{z_1, \ldots, z_m\}. \qquad (3)$$

Then r is the rank of M if and only if it is the rank of

$$\begin{pmatrix} I & -AE & 0 \\ 0 & I & 0 \\ 0 & -GE & I_m \end{pmatrix} M = \begin{pmatrix} 0 & B' - AEC & G' - AEH \\ B & C & H' \\ 0 & H - GEC & K - GEH' \end{pmatrix}, \qquad (4)$$

where I is the r-rowed identity matrix and I_m is the m-rowed identity matrix. This yields the

THEOREM. *The factor matrix M has rank r if and only if X, Y, Z are chosen so that*

$$K = GEH', \qquad H = GEC, \qquad G' = AEH', \qquad B' = AEC. \tag{5}$$

We observe that our hypothesis that r is the ideal rank implies that r is the rank of the $(r + 1)$-rowed square matrix

$$\begin{pmatrix} B & h_j' \\ g_i & k_{ij} \end{pmatrix}$$

obtained by using the i-th row g_i of G, the j-th column h_j' of H', and the element k_{ij} in the i-th row and j-th column of K_0 for $i \neq j$. Then r is the rank of

$$\begin{pmatrix} I & 0 \\ -g_iE & 1 \end{pmatrix}\begin{pmatrix} B & h_j' \\ g_i & k_{ij} \end{pmatrix} = \begin{pmatrix} B & h_j' \\ 0 & k_{ij} - g_iEh_j' \end{pmatrix},$$

that is, $k_{ij} = g_iEh_j'$. But then

$$GEH' - K_0$$

is a diagonal matrix and we satisfy the first relation of (5) by taking $Z = GEH' - K_0$. The second and third relations may be written as

$$GEY = H - GEC_0, \qquad XEH' = G' - A_0EH', \tag{6}$$

and we have the basis for the following:

DETERMINATION OF DIAGONAL ELEMENTS

I. *Estimate the ideal rank r.* This results in the fixing of our notation for B, A_0, C_0, G, H, K_0.

II. *Find the inverse E of B.* If our estimate of r is too great it will not be possible to carry this out.

III. *Compute $GE = (\sigma_{ij})$ where $i = 1, \ldots, m$ and $j = 1, \ldots, r$.*

IV. *Compute $(GE)H'$.* The nondiagonal elements of this m by n product will coincide with the corresponding nondiagonal elements of K_0 unless our estimate of r is too small. If they do not coincide the choice of B needs to be modified and it is a simple matter to complete the determination of the inverse of the new B. If they do coincide the diagonal elements of $(GE)H'$ are the desired communalities $z_1, z_2, \ldots, z_m$.

V. *Compute, for every nonzero element σ_{ij} of GE, the element in the i-th row and j-th column of $(GE)C_0$. Subtract this computed element from the correspondingly placed element of H and divide by σ_{ij}.* The result is y_j, the $(r + j)$-th diagonal element of M. In most cases this will yield m determinations of y_j and a weighted mean obtained by summing elements in columns would probably give a best value.

VI. *Compute $EH' = (\lambda_{ji})$, where $j = 1, \ldots, r$ and $i = 1, \ldots, m$. Find, for every nonzero λ_{ji}, the element in the j-th row and i-th column of $A_0(EH')$*

Subtract the result from the correspondingly placed element in G' and divide by λ_{ji}. The result will be x_j, the j-th diagonal element of M.

In the cases most usually encountered the procedure above completely determines all diagonal elements and gives a full solution of our problem. Note that if we determine X so that $G' = AEH'$ as in step VI, then $H - GEC = H - HE'AEC = HE'(B' - AEC)$. If H has rank at least r it will follow that the determination of Y as in step V, which yields $H = GEC$, will imply that $B' = AEC$, M has rank r. Conversely our remaining work in which we investigate $B' = AEC$ implies that if we make $B' = AEC$ and $G' = AEH'$ then we will have $H = GEC$.

It is necessary to investigate the relation $B' = AEC$ since we might have a case where $n = 2r$, G, H, and K_0 do not appear, and the procedure outlined above is vacuous. In this case we proceed to compute $E = (\delta_{ij})$ as usual, and in all cases where the procedure above does not determine X and Y, we continue as follows:

VII. *Compute the matrix products*

$$(\alpha_{ij}) = -EC_0, \quad (\beta_{ij}) = B' - A_0(EC_0), (\gamma_{ij}) = A_0E \quad (i, j = 1, \ldots, r). \quad (7)$$

VIII. *Solve the equations*

$$x_k[x_i(\alpha_{ij}\gamma_{kj} - \gamma_{ij}\alpha_{kj}) + (\beta_{ij}\gamma_{kj} - \delta_{ij}\alpha_{kj})] = x_i(\gamma_{ij}\beta_{kj} - \alpha_{ij}\delta_{kj}) + (\delta_{ij}\beta_{kj} - \beta_{ij}\delta_{kj}) \quad (i \neq k;\ i, j, k = 1, \ldots, r). \quad (8)$$

These equations may be used normally to solve for $r - 1$ of the unknown diagonal elements x_k as quotients of two linear functions of the remaining one. The consistency of our system will then usually determine the remaining unknown. In all cases the solution must be such that

$$|A| \neq 0. \quad (9)$$

Observe that in the exceptional case where $A_0 = C_0 = 0$ and $B = I$ the diagonal elements of A are any nonzero numbers and $C = A^{-1}$, $Y = X^{-1}$.

IX. *Solve the equations*

$$(x_i\gamma_{ij} + \delta_{ij})y_j = x_i\alpha_{ij} + \beta_{ij} \quad (i, j = 1, \ldots, r) \quad (10)$$

for the diagonal elements y_j.

Equations (10) normally yield r determinations of each y_i. These equations may also be used as checks if the diagonal elements have been determined by the use of V, VI.

Tests VIII and IX are equivalent to $B' = AEC$. For B is nonsingular and this matric equation implies (9). Then $B' = AEC(A_0 + X)E(C_0 + Y)$, $(B' - A_0EC_0) + X(-EC_0) = [(A_0 + X)E]Y$. This equation is equivalent to (10). For each value of j the equations (10) are a system of r equations in a single unknown y_j with matrix Γ_j a column of the nonsingular matrix AE, and Γ_j has rank one. Consistency of (10) demands that the rank of the augmented matrices must be one, that is, for all $i \neq k$ and j we must have

$$(x_i\alpha_{ij} + \beta_{ij})(x_k\gamma_{kj} + \delta_{kj}) = (x_i\gamma_{ij} + \delta_{ij})(x_k\alpha_{kj} + \beta_{kj}). \tag{11}$$

These equations are another form of (8).

The various determinations of the same diagonal elements may be employed to determine them accurately in cases where there are experimental errors in the given correlation coefficients. It should be possible then to obtain data indicating clearly the correct values, and final corroboration will be obtained by factorization of the matrix M.

EXAMPLE

$$M_0 = \begin{pmatrix} 0 & 0.72 & -0.61 & 0.92 & -0.26 & 0.59 \\ 0.72 & 0 & -0.80 & 0.68 & 0.20 & 0.68 \\ -0.61 & -0.80 & 0 & -0.54 & -0.40 & -0.69 \\ 0.92 & 0.68 & -0.54 & 0 & -0.39 & 0.55 \\ -0.26 & 0.20 & -0.40 & -0.39 & 0 & 0.21 \\ 0.59 & 0.68 & -0.69 & 0.55 & 0.21 & 0 \end{pmatrix}.$$

Then

$$B = \begin{pmatrix} -0.61 & -0.80 \\ 0.92 & 0.68 \end{pmatrix}, \qquad E = \begin{pmatrix} 2.118 & 2.492 \\ -2.866 & -1.900 \end{pmatrix},$$

so that

$$GE = \begin{pmatrix} -1.124 & -1.028 \\ -0.699 & 0.178 \end{pmatrix}, \qquad GEH' = \begin{pmatrix} 0.851 & 0.210 \\ 0.210 & 0.58 \end{pmatrix}$$

This verifies our assumption that the ideal rank is two. We also compute

$$GEC_0 = \begin{pmatrix} -1.124 & -1.028 \\ -0.699 & 0.178 \end{pmatrix}\begin{pmatrix} 0 & -0.54 \\ -0.54 & 0 \end{pmatrix} = \begin{pmatrix} 0.555 & 0.607 \\ -0.096 & 0.377 \end{pmatrix}, \qquad H - GEC_0 = \begin{pmatrix} -0.955 & -0.997 \\ -0.544 & 0.173 \end{pmatrix},$$

to give $y_1 = 0.850$, $y_1 = 0.850$, $y_2 = 0.970$, $y_2 = 0.972$. Similarly

$$EH' = \begin{pmatrix} -1.819 & -0.091 \\ 1.887 & 0.933 \end{pmatrix}, \qquad A_0EH' = \begin{pmatrix} -1.359 & 0.672 \\ -1.310 & -0.066 \end{pmatrix},$$

$$G' - A_0(EH') = \begin{pmatrix} -1.619 & -0.082 \\ 1.510 & 0.746 \end{pmatrix},$$

so that $x_1 = 0.890$, $x_1 = 0.901$, $x_2 = 0.800$, $x_2 = 0.800$. The bad value of 0.901 is due to the smallness of the numbers 0.082 and 0.091, and the weighted mean

$$\frac{1.619 + 0.082}{1.819 + 0.091} = 0.891$$

indicates that the correct value is 0.89. Actually $M = FF'$, where

$$F' = \begin{pmatrix} 0.8 & 0.4 & -0.2 & 0.9 & -0.7 & 0.3 \\ 0.5 & 0.8 & -0.9 & 0.4 & 0.6 & 0.7 \end{pmatrix}.$$

[1] Cf. Thurstone, L. L., *The Vectors of Mind*, Chicago, 1935, and Lederman, W., "On a Problem Concerning Matrices with Variable Diagonal Elements," *Proc. Roy. Soc. Edin.*, Vol. 60, (1939), 17 pp.

[2] Actually $M = P\Lambda P'$ where Λ is a diagonal matrix with its last $n - r$ diagonal elements zero, P is an orthogonal matrix. Then F is the product of the first r columns of P by the respective square roots of the first r diagonal elements of Λ, that is, the square roots of the characteristic roots of M.

LAPLACE TRANSFORMS OF A CLASS OF HIGHER DIMENSIONAL VARIETIES IN A PROJECTIVE SPACE OF n DIMENSIONS

By Shiing-shen Chern

Institute for Advanced Study

Communicated February 15, 1944

The theory of conjugate nets and their Laplace transforms has been the object of extensive study. Its extension to r-dimensional varieties in a projective space of n dimensions, $r < n$, has so far not been very successful. The object of this note is to indicate a possible way that the theory can be extended.

Consider a projective space of n dimensions. To develop the differential geometry in the space we employ Cartan's method of moving frames, taking as projective frame the set of $n + 1$ linearly independent vectors $A, A_1, \ldots, A_n$ (each with $n + 1$ components), determined up to a common factor. The equations of infinitesimal displacement are of the form

$$\left.\begin{aligned} dA &= \omega_{00}A + \omega_1 A_1 + \ldots + \omega_n A_n \\ dA_1 &= \omega_{10}A + \omega_{11}A_1 + \ldots + \omega_{1n}A_n \\ &\vdots \\ dA_n &= \omega_{n0}A + \omega_{n1}A_1 + \ldots + \omega_{nn}A_n \end{aligned}\right\} \quad (1)$$

where the ω's are Pfaffian forms in the parameters on which depend the frames. By remarking that the exterior derivatives of the left-hand sides of (1) are zero, we get the so-called equations of structure of the space:

$$\left.\begin{aligned} \omega'_{00} &= [\omega_1\omega_{10}] + \ldots + [\omega_n\omega_{n0}] \\ \omega'_i &= [\omega_{00}\omega_i] + [\omega_1\omega_{1i}] + \ldots + [\omega_n\omega_{ni}] \\ \omega'_{ij} &= [\omega_{i0}\omega_j] + [\omega_{i1}\omega_{1j}] + \ldots + [\omega_{in}\omega_{nj}],\ i, j = 1, \ldots, n \\ \omega'_{i0} &= [\omega_{i0}\omega_{00}] + [\omega_{i1}\omega_{10}] + \ldots + [\omega_{in}\omega_{n0}] \end{aligned}\right\} \quad (2)$$

Let V_r be a variety of r dimensions in the space. We make the convention that a linear combination of the vectors of a frame represents a point whose homogeneous coördinates are its components. To each point of V_r we attach a frame such that A coincides with the point, $A_1, \ldots, A_r$ lie in the tangent r-plane, and $A_{r+1}, \ldots, A_n$ $n - r$ other points of the space. For such a family of frames over V_r we have

$$\left\{\begin{aligned} \omega_{r+1} &= \ldots = \omega_n = 0, \\ \omega_{i\alpha} &= \sum_{j=1}^{r} g_{ij\alpha}\omega_j,\ i = 1, \ldots, r,\ \alpha = r + 1, \ldots, n \end{aligned}\right. \quad (3)$$

where $g_{ij\alpha}$ are functions of parameters on the variety and are symmetric with respect to i, j. Construct the quadratic forms

$$\Phi_\alpha = \omega_1\omega_{1\alpha} + \ldots + \omega_r\omega_{r\alpha},\ \alpha = r + 1, \ldots, n \quad (4)$$

and suppose q of them be linearly independent. This number q is an arithmetic invariant. We can choose the frame attached to the point A such that $\Phi_{r+1}, \ldots, \Phi_{r+q}$ are linearly independent and $\Phi_{r+q+1}, \ldots, \Phi_n$ are identically zero. We shall call the net of quadric cones of vertex A defined by the equation

$$\lambda_{r+1}\Phi_{r+1} + \ldots + \lambda_{r+q}\Phi_{r+q} = 0 \quad (5)$$

the asymptotic net at A. It generalizes the asymptotic tangents of a surface.

E. Cartan[1] studied the varieties of r dimensions whose asymptotic net is reducible to the form

$$\lambda_1\omega_1^2 + \ldots + \lambda_r\omega_r^2 = 0 \quad (6)$$

Cartan proved that such varieties depend on $r(r - 1)$ arbitrary functions in two variables. It is not difficult to prove that for $r = 2$ these varieties are in general the surfaces sustaining conjugate nets.

From the last remark we can deduce a notable geometrical property of these varieties. Suppose the frames attached to the points of V_r are so chosen that the asymptotic net has the equation (6). Then the lines AA_i, $i = 1, \ldots, r$, are invariantly related to the variety. They define on the variety r families of curves such that through every point passes one and only one curve of each family. If the tangent r-plane is displaced along a curve of the ith family, its intersection with a neighboring tangent

r-plane is the tangent $(r - 1)$-plane $AA_1 \ldots A_{i-1}A_{i+1} \ldots A_r$. This is clearly a generalization of the well-known property of conjugate nets.

It turns out that this property is characteristic to the varieties under consideration. In fact, suppose that a variety of r dimensions possesses r families of curves such that when the tangent r-plane is displaced along a curve of one family, its intersection with a neighboring tangent r-plane is the tangent $(r - 1)$-plane tangent to the other $r - 1$ curves. Then, under a certain generality assumption, the variety has the property that its asymptotic net is reducible to the equation (6). The generality assumption has to be made, because there are some degenerate cases, which are not interesting. A simple and reasonable assumption of this kind is that the osculating planes of the variety are of $2r$ dimensions.

The above results can be described as a geometrical interpretation of the varieties studied by Cartan. To define for such varieties a kind of transformation generalizing the transformation of Laplace, we proceed as follows:

Let V_r be such a variety whose r families of curves have at the point A the tangents AA_i, $i = 1, \ldots, r$. Then the following theorem can be proved: On each tangent AA_i there exist $r - 1$ points A_{ij} $(j \neq i)$ such that when the s-plane $AA_{i_1} \ldots A_{i_s}$ is displaced along AA_j, $j \neq i_1, \ldots, i_s$, its intersection with a neighboring s-plane is the $(s - 1)$-plane $A_{i_1j} \ldots A_{i_sj}$.

Associated to each point A we have therefore $r(r - 1)$ points A_{ij}. When A describes our variety, each point A_{ij} describes a variety of dimension $\leqq r$. It can be proved that if the locus of A_{ij} is of dimension r, it also has the property described above. It is natural to call these varieties the Laplace transforms of V_r, there being altogether $r(r - 1)$ Laplace transforms. The Laplace transforms of a Laplace transform of V_r will consist of V_r, some of the Laplace transforms of V_r, and some new varieties. There is a great number of interesting problems which generalize the classical theory of periodic Laplace sequences. The problems seem to be particularly fascinating, because it is probable to expect some results which are not direct generalizations of the theory of surfaces.

The proofs of the above results offer no difficulty and are therefore omitted. For long time there was the opinion that the projective differential geometry of r-dimensional varieties in a space of n dimensions has serious analytical difficulties. We wish to point out here that Cartan's method of moving frames offers a workable scheme which, at least theoretically, can push to an end the problems which are of geometrical interest.

---

[1] Cartan, E., "Sur les variétés de courbure constante d'un espace euclidien ou non-euclidien," *Bull. Soc. math. France*, **47**, 125–160 (1919); **48**, 132–208 (1920).

PROCEEDINGS
OF THE
NATIONAL ACADEMY OF SCIENCES
Volume 30 May 15, 1944 Number 5
Copyright 1944 by the National Academy of Sciences

STRAIN SPECIFICITY AND PRODUCTION OF ANTIBIOTIC SUBSTANCES. III. PENICILLIUM NOTATUM—CHRYSOGENUM GROUP*

By Selman A. Waksman and H. Christine Reilly

New Jersey Agricultural Experiment Station, Rutgers University, New Brunswick, N. J.

Communicated April 7, 1944

It has previously been shown[15,20] that although certain groups of microorganisms are characterized by the capacity to produce antibiotic substances, great variation may be found in the quantitative production of the specific substance among the strains within each group. Aside from the natural variation of a group of organisms to bring about a certain physiological reaction, this phenomenon is markedly influenced by three environmental and nutritional factors: (1) the composition of the medium, especially the presence of peptones, carbohydrates and certain growth-promoting or other nutritional substances;[17] (2) conditions of growth, notably aeration and temperature of incubation—some strains grow better in stationary cultures, where they usually produce a surface pellicle, whereas others grow best in submerged and agitated cultures, in the form of suspended colonies or masses of cell material throughout the medium; and (3) changes in the composition of the medium produced by different strains of the same organism—some strains of *Aspergillus clavatus*, for example, change the reaction of the medium to alkaline, rather than to acid, thus favoring the rapid destruction of the antibiotic substance clavacin.[20]

The fact that many microörganisms produce more than one antibiotic substance and the fact that the activity of these substances is measured by a single biological method, based upon their selective antibacterial activity in artificial culture media against a single test organism, often tend to introduce further confusion in the interpretation of the results obtained. The following two illustrations will suffice: (1) The *Penicillium notatum-chrysogenum* group of fungi produce penicillin, a relatively non-toxic compound highly active against gram-positive bacteria, and a protein that functions as a glucose-oxidase and that exerts an antibacterial action only in the presence of glucose and other carbohydrates. This often tended to confuse the isolation and characterization of penicillin, especially when media used for the testing of the antibiotic substance contained glucose.[8,5,13] (2) The *A. fumigatus* group of fungi produce, in addition to the pigment fumigatin and the non-nitrogenous and sulfur-free fumigacin,[18] also the nitrogen and sulfur-containing gliotoxin.[10,16] The presence[18] or absence[3] of a small amount of gliotoxin admixed with the fumigacin crystals led to its characterization as a different compound.

Another series of complications in the study of antibiotic substances arose from the fact that the same substance may be produced by different fungi. Thus, the designation of a single compound as clavacin, claviformin, patulin and clavatin is based either upon the insufficient characterization of the crude *vs.* the crystalline compounds or upon its production by different types of fungi.[12,7,14]. It is now fairly well established that penicillin or penicillin-like compounds are produced by a large number of different fungi, belonging not only to different species but even to different genera. In spite of this, the substance formed by members of the *A. flavus* group and by different species of *Penicillium* has not only been given different designations (flavicin,[2] flavacidin,[9] gigantic acid[11], parasiticin[4a]) but synonyms have already been piled upon synonyms, with such additional ones as aspergillin and flavatin.

Some of the groups of fungi-producing antibiotic substances show extreme variability in this respect; here belong *A. clavatus* and *A. flavus-oryzae*, discussed previously,[19,15] and *P. notatum-chrysogenum*, the subject of this paper. Different strains obtained from different substrates and, in many instances, isolated from the same mother culture show distinct physiological differences that influence not only the abundance, but often the very nature, of the antibiotic substances produced.[4,6]

Strains Used.—In the following investigations on the production of penicillin and penicillin-like substances by different fungi, 17 cultures were used. These comprised 13 strains of *P. notatum* obtained from different sources, 3 other species of *Penicillium* belonging to the group of green Penicillia, and one species of *Aspergillus*. The numbers and sources of these fungi are briefly summarized:

No. 40.	*P. notatum*, received from Dr. C. Thom of Washington, D. C., in 1936, as No. 144-5112.1.
No. 41.	*P. notatum*, Fleming strain, received from Dr. C. Thom in December, 1939.
No. 42.	*P. notatum*, Westling strain, received from Dr. R. St. John-Brooks of London, in December, 1940, who in turn received it through Dr. H. Raistrick, from C. B. S. Baarn Collection in Holland, in 1925.
No. 43.	*P. notatum*, received from Dr. R. St. John-Brooks, in December, 1940, who in turn received it from Dr. Fleming in 1934.
Nos. 44, 45.	*P. notatum*, isolated by Prof. A. Callenger of Leeds University, from bread, in 1938, identified by Dr. C. Thom; received from Dr. R. St. John-Brooks in 1940,

No. 46.	*P. notatum*, Reid's culture of Fleming's strain, received from Dr. Reid in July, 1941.
No. 47.	*P. notatum*, Heatley's culture of Fleming's strain, received from Dr. Heatley in December, 1941.
No. 48.	*P. janthinellum*, isolated from the soil and kept for several years in culture collection.
No. 50.	*P. notatum*, NRRL No. 832; represents a strain isolated by Dr. Biourge in Belgium, and received from him by C. Thom; received from Dr. K. B. Raper in February, 1943.
No. 51.	*P. notatum*, isolated at the NRRL from the Squibb Strain, No. 1249, B21; received from Dr. K. B. Raper in March, 1943.
No. 136.	*P. spinulosum*, isolated by Dr. M. B. Morrow of Austin, Texas, from a Guatemala soil; received in August, 1942.
No. 137.	*A. flavus*, isolated from the soil in 1941, and found[16] to be capable of producing a penicillin-like substance under certain conditions of culture.
No. 161.	*Penicillium* sp. a green Penicillium isolated from the soil in 1941.[17]
No. 7813.	*P. notatum;* Dr. C. Thom's culture No. 144B69, received from ATCC, in August, 1943 (NRRL No. 832).
No. 8251.	*P. notatum*, Fleming strain, Thom's culture No. 144-5767, received from ATCC, in August, 1943.
No. 8537.	*P. notatum* received by the NRRL from Dr. Florey; this is presumably the same culture as No. 47; it was received from the ATCC in August, 1943.

Conditions of Culture.—Three different methods were used for the cultivation of these fungi and for the production of penicillin; namely, the liquid stationary, the liquid submerged, and the deep agar-tube methods. In the first, the medium, consisting of a modified Czapek-Dox broth, with brown sugar in place of glucose, and supplemented with corn steep liquor[16,17] and $CaCO_3$, was placed in Erlenmeyer flasks to a depth of about 1 to 1.5 cm.; the flasks were incubated at about 22°C. for varying periods of time. In the submerged culture, the medium was placed, in 100-ml. portions in 250-ml. flasks, in a shaking machine, and incubated at 28°C. In the third method,[19] a medium similar to the above, but supplemented with 2 per cent agar, was placed in 20-ml. portions in test tubes, measuring 18 × 140 mm., and the tubes were inoculated at the surface.

Methods of Testing.—Three different methods were used for testing the antibiotic activity of the cultures: (1) the dilution method, using the streak agar plate,[21] with *Staphylococcus aureus* or *Bacillus subtilis* as test organism; (2) the agar diffusion or cup method with *B. subtilis* as the test organism;[1] (3) the agar tube method,[19] with *B. subtilis* as the test organism.

Results Obtained.—In the first experiment, the effect of corn steep upon the production of penicillin was determined, using strain No. 50, by growing the culture both in a stationary and a submerged condition. The results (table 1) gave the highest activity for the submerged cultures containing corn steep, a fact already established by workers in the Northern Regional Laboratory. A more detailed study of the production of penicillin under submerged and stationary cultures gave a maximum activity in 4 days for the submerged and in 10 to 12 days for the stationary cultures.

A comparison of the production of penicillin by 15 different cultures of fungi, grown under both stationary and submerged conditions (table 2), brought out marked differences in activity. Some of the cultures were found to be highly active under both submerged and stationary conditions (No. 7813); some were inactive or had only limited activity under both conditions (Nos. 41, 43, 44, 45, 48); some gave a higher activity under one set of conditions and others under another. On comparing the course of production of penicillin by *P. notatum* (No. 50) and *A. flavus* (No. 137), both grown in a submerged and agitated state, it was found that although the activity of the first reached a higher maximum, it tended to disappear more rapidly than that of the second. Bacteriostatic spectra of the two culture filtrates were made after 5 days' incubation, when the activity of both was at a maximum, and found to be identical, corresponding to that of isolated penicillin.

TABLE 1

PENICILLIN PRODUCTION BY *P. notatum*, AS INFLUENCED BY ADDITION OF CORN STEEP LIQUOR AND BY SHAKING OF CULTURES

MEDIUM	GROWTH OF CULTURES	INCUBATION, DAYS	PLATE METHOD, DILUTION UNITS[a]	CUP METHOD, OXFORD UNITS[b]	pH
BS[c]	Submerged	3	30	0.9	6.7
BS	Submerged	5	100	2.3	8.2
BS	Submerged	7	30	6.3	8.8
BS	Stationary	5	100	0.3	6.0
BS	Stationary	7	<30	1.8	7.2
CSM	Submerged	2	300	5.0	6.5
CSM	Submerged	3	1000	5.4	7.9
CSM	Submerged	5	1000	15.4	8.2
CSM	Submerged	7	300	11.8	8.8
CSM	Stationary	4	<30	1.6	...
CSM	Stationary	5	100	0.5	6.6
CSM	Stationary	7	30	3.7	6.8

[a] *B. subtilis* as test organism.

[b] *Staph. aureus* as test organism.

[c] BS = Brown sugar medium, containing 4 per cent brown sugar and 2 p. p. m. $ZnSO_4 \cdot 7H_2O$, in addition to other salts; CSM = corn steep medium, containing 2 per cent brown sugar, 1 per cent $CaCO_3$ and 1.5 per cent corn steep, in addition to other salts.

The usefulness of the agar tube method for making rapid comparisons of the production of penicillin by different fungi, as well as for studying the effect of the composition of medium and conditions of growth upon this reaction, is illustrated in table 3. When the Czapek-Dox medium was used, No. 136 showed no activity at all; however, in brown-sugar media it proved to be as efficient as two of the best *P. notatum* strains (Nos. 41, 50). On the other hand, No. 161, which was nearly as good as the *P. notatum*

TABLE 2

PRODUCTION OF PENICILLIN BY DIFFERENT STRAINS OF *P. notatum* AND OTHER FUNGI

CULTURE NO.	SUBMERGED CULTURES		STATIONARY CULTURES	
	INCUBATION, DAYS	ACTIVITY, DILUTION UNITS[a]	INCUBATION, DAYS	ACTIVITY, DILUTION UNITS
40	5	50	7	50
40	7	300	..	..
41	7	0	9	0
42	4	20	7	100
42	6	50	9	100
43	7	5	10	..
44	3	0	7	10
44	7	0	..	..
45	4	0	7	10
45	7	0	..	..
46	5	30	9	10
46	7	30	..	..
47	7	50	10	..
47	9	100	..	..
48	5	0	7	0
50	3	150	7	75
50	5	300	..	..
51	7	300	..	..
136	4	300	6	75
136	6	10	..	..
7813	4	1000	7	500
7813	6	1000	9	300
7813	8	1500	..	..
8251	4	30	7	30
8251	6	30	10	30
8537	4	30	7	30
8537	6	100	9	30

[a] *Staph. aureus* used as test organism.

TABLE 3

INFLUENCE OF COMPOSITION OF MEDIUM UPON THE PRODUCTION OF PENICILLIN BY DIFFERENT ORGANISMS, AS MEASURED BY AGAR-TUBE METHOD

5-Day Incubation, Top Layer, Mm. of Zone Produced on *B. subtilis* Inoculated Plate

ORGANISM NO.	COMPOSITION OF MEDIUM			
	CZAPEK-DOX	BROWN SUGAR[a]	CORN STEEP[b]	ZINC[c]
41	19.0	21.0	40.0	41.0
50	19.5	23.0	39.5	39.5
136	0	23.0	36.0	38.0
161	18.0	19.0	26.0	26.0

[a] Brown sugar in place of glucose.

[b] Supplementary to above, 1.5 per cent.

[c] $ZnSO_4.7H_2O$, 0.001 per cent, supplementary to above.

strains when grown in the synthetic medium, proved much inferior in the corn steep media.

A comparison of the production and diffusion of penicillin by a larger number of cultures of fungi by the agar tube method is recorded in table 4. Some of the strains that gave rather inferior results in liquid media proved to be excellent producers of penicillin in the agar medium. Number 41, for example, gave virtually no activity in either submerged or stationary cultures, as shown in table 2, but gave excellent activity in the agar culture. Number 46 that gave only low activity in liquid media gave as high activity as the best strain, 7813. These results thus serve to emphasize that the strain of the organism, the conditions of growth and the composition of the medium are the three essential factors in the production of penicillin.

Discussion.—The production of penicillin is not limited to any one species of fungi or even to the genus *Penicillium*. Different species and different strains of the same species appear to produce the maximum amount of penicillin or penicillin-like substances under different conditions of culture. Whereas some strains are most active under one set of conditions, other strains exert their greatest activity under other conditions. The fact that, in addition to penicillin, a second antibiotic factor is produced by most of the penicillin-producing fungi further complicates this relationship; this factor may be active only in the presence of glucose, as in the case of the *P. notatum* strains, or its activity may be independent of the glucose in the medium, as in the case of the aspergillic acid produced by *A. flavus*. The production of the second factor is also influenced by the composition of the medium and by conditions of growth.

TABLE 4

PENICILLIN PRODUCTION BY DIFFERENT STRAINS OF *P. notatum* AND OTHER FUNGI
Agar-Tube Method, Incubation 6 Days, Mm. of Zone Produced on *B. subtilis* Plate

ORGANISM NO.	TOP OF COLUMN	MIDDLE OF COLUMN	BOTTOM OF COLUMN
40	40.0	32.0	19.5
41	42.5	26.5	0
42	36.3	23.5	0
44	34.8	27.5	0
45	30.8	12.5	0
46	43.0	32.0	16.3
47	40.0	29.8	0
50	43.5	30.5	11.0
136	39.0	31.5	0
161	25.5	11.5	0
7813	44.0	35.8	18.0
8251	39.5	24.5	0
8537	38.5	25.8	0

---

[*] Journal Series Paper, New Jersey Agricultural Experiment Station, Rutgers University, Department of Soil Microbiology.

[1] Abraham, E. P., Chain, E., Fletcher, C. M., Gardner, A. D., Heatley, N. G.,

Jennings, M. A., and Florey, H. W., *Lancet*, **241**, 177–189; *Nature*, **148**, 758–759; **149**, 356 (1941–1942).

[2] Bush, M. T., and Goth, A., *Jour. Pharmacol.*, **78**, 164–169 (1943).

[3] Chain, E., Florey, H. W., Jennings, M. A., and Williams, T. I., *Brit. Jour. Exp. Path.*, **24**, 108–119; see also Crowfoot, D. M., and Low, B. W., *Ibid.*, **24**, 120 (1943).

[4] Clutterbuck, P. W., Lovell, R., and Raistrick, H., *Biochem. Jour.*, **26**, 1907–1918 (1932).

[4a] Cook, A. H., and Lacey, M. S., *Nature*, **153**, 460 (1944).

[5] Coulthard, C. E., Short, W. F., Skrimshire, G. H. E., Birkinshaw, J. H., Michaelis, R., Sykes, G., Standfast, A. F. B., and Raistrick, H., *Nature*, **150**, 634–645; *Jour. Biol. Chem.*, **148**, 459–460 (1942).

[6] Foster, J. W., Woodruff, H. B., and McDaniel, L. E., *Jour. Bact.*, **46**, 421–433 (1943).

[7] Hooper, I. R., Anderson, H. W., Skell, P., and Carter, H. E., *Science*, **99**, 16 (1944).

[8] Kocholaty, W., *Science*, **97**, 186–187; *Arch. Biochem.*, **2**, 73–86 (1943).

[9] McKee, C. M., Rake, G., and Houck, C. L., *Jour. Bact.*, **47**, 187–197 (1944).

[10] Menzel, A. E. O., Wintersteiner, O., and Hoogerheide, J. C., *Jour. Biol. Chem.*, **152**, 419–429 (1944).

[11] Philpot, F. J., *Nature*, **152**, 725 (1943).

[12] Raistrick, H., Birkinshaw, J. H., Michael, S. E., Bracken, A., Gye, W. E., and Hopkins, W. A., *Lancet*, **245**, 625–635; see also *Ibid.*, **245**, 684 (1943).

[13] Roberts, E. C., Cain, C. K., Muir, R. D., Reithel, F. J., Gaby, W. L., VanBruggen, J. T., Homan, D. M., Katzman, P. A., Jones, L. R., and Doisy, E. A., *Jour. Biol. Chem.*, **147**, 47–58; **148**, 365–378 (1943).

[14] Waksman, S. A., *Science*, **99**, 220–221 (1944).

[15] Waksman, S. A., and Bugie, E., *Proc. Nat. Acad. Sci.*, **29**, 282–288 (1943).

[16] Waksman, S. A., and Geiger, W. B., *Jour. Bact.*, **47**, 391–397 (1944).

[17] Waksman, S. A., and Horning, E., *Mycologia*, **35**, 47–65 (1943).

[18] Waksman, S. A., Horning, E., and Spencer, E. L., *Science*, **96**, 202–203 (1942).

[19] Waksman, S. A., and Reilly, H. C., *Jour. Bact.*, **47**, 308–309 (1944).

[20] Waksman, S. A., and Schatz, A., *Proc. Nat. Acad. Sci.*, **29**, 74–79 (1943).

[21] Waksman, S. A., and Woodruff, H. B., *Jour. Bact.*, **40**, 581–600 (1940).

PERMUTABLE PERIODIC TRANSFORMATIONS

BY P. A. SMITH

DEPARTMENT OF MATHEMATICS, COLUMBIA UNIVERSITY

Communicated April 11, 1944

The so-called "special homologies" were introduced as useful tools[1] in the topology of finite cyclic transformation groups. Up to the present they have found no application in connection with non-cyclic groups. The following remarks indicate, however, that such applications do exist.

Let K be a finite simplicial complex and let $\mathfrak{T} = \{ t \}$ be a group of regular[2] simplicial homeomorphisms of K into itself. Let G be an additive coefficient group. Polynomials in elements of $\mathfrak{T}$ with integer coefficients can be considered as operators on the chains of K, equally well as operators on the homology classes of K since they permute with the boundary operator F. The totality of these operators is a ring $P = \{\rho, \sigma, \ldots\}$ which is commutative when $\mathfrak{T}$ is. We consider only the commutative case in the present note.

We denote by 0 the operator which converts all chains into null chains. Let p be the order of an element t of $\mathfrak{T}$. We denote either one of the operators

$$1 - t, \qquad 1 + t + \ldots + t^{p-1}$$

by $\rho(t)$, the other by $\bar{\rho}(t)$. These elements are zero-divisors of P since $\rho\bar{\rho} = 1 - t^p = 0$.

Given a set of elements $t_1, \ldots, t_m$ of $\mathfrak{T}$, we define a set $I(t_1, \ldots, t_m)$ of simplexes of K by the rule that E is in I if and only if E is invariant under some transformation other than the identity in the subgroup generated by $t_1, \ldots, t_m$.

LEMMA 1. *Let $t_1, \ldots, t_m$ be independent[3] elements of $\mathfrak{T}$ and let X be a chain in $K - I(t_1, \ldots, t_m)$. A necessary and sufficient condition that*

$$\rho(t_1)\rho(t_2) \ldots \rho(t_m)X = 0$$

is that X be expressible in the form

$$X = \bar{\rho}(t_1)Y_1 + \ldots + \bar{\rho}(t_m)Y_m,$$

the Y's being chains in $K - I$.

LEMMA 2. *Let the t's and X be as in lemma 1. A necessary and sufficient condition that*

$$\rho(t_1)X = \ldots = \rho(t_m)X = 0$$

is that X be expressible in the form

$$X = \bar{\rho}(t_1)\bar{\rho}(t_2) \ldots \bar{\rho}(t_m)Z, \qquad Z \subset K - I.$$

A proof of lemma 1 for the case in which $m = 1$ has been given elsewhere[4] and its generalization to arbitrary M is straightforward. Lemma 2 is a fairly direct corollary of lemma 1.

Application. Let p be a prime and let M be a homological n-sphere over the group $\mathfrak{J}_p$ of integers modulo p. Let t be a transformation of period p operating in M. The fixed-point set L of t is[5] a homological r-sphere over $\mathfrak{J}_p$ ($r \leqq n$), possibly null ($r = -1$). If M satisfies certain local conditions, for example if M is locally euclidean, r is necessarily less than n. We assume here that M satisfies such conditions.

THEOREM. *If t and s are permutable transformations of prime period p operating in M and if their fixed-point sets are identical (possibly null), then s is a power of t.*

COROLLARY.[6] *There cannot operate in M a commutative group of fixed-point-free transformations which is of order p^a, $a > 1$, and of type $(p, \ldots, p)$.*

"Fixedpoint-free" refers, of course, to transformations other than the identity.

COROLLARY. *M cannot admit a commutative transformation group of order p^k and type $(p, \ldots, p)$ unless $k \leqq (n + 1)/2$.*

We shall prove the theorem for the regular simplicial case and refer elsewhere[7] for the method of passing from the simplicial to the topological case.

We suppose then that M is a finite n-complex and that $I(s) = I(t) = L$, say. L is a homological r-sphere over $\mathfrak{Z}_p$. We suppose that s is not a power of t. Then s and t generate a group of order p and type (p, p).

Let $\rho_n, \rho_{n-1}, \ldots$ stand alternately for $1 - t$ and $1 + t + \ldots + t^{p-1}$ and let $\delta = 1 - s$, $\sigma = 1 + s + \ldots + s^{p-1}$. Let $\mathfrak{H}^h_p$ be the group of ρ_h-homology[8] classes of (M, t). There exists for each h a decomposition[9]

$$\mathfrak{H}^h_p = \mathfrak{O}^h_p \times \mathfrak{H}^h_{0p} \tag{1}$$

such that the cycles of $\mathfrak{O}^h_p$ are in $M - L$, those of $\mathfrak{H}^h_{0p}$ in L. Under the present hypotheses *the groups $\mathfrak{O}^n_p, \ldots, \mathfrak{O}^{r+1}_p$ are cyclic[10] of order p* and

$$\mathfrak{H}^h_p = \mathfrak{O}^h_p, \qquad \mathfrak{H}^h_{0p} = 0, \qquad (h = n, \ldots, r + 1) \tag{2}$$

$$\mathfrak{H}^r_p = \mathfrak{H}^r_{0p}, \qquad \mathfrak{O}^r_p = 0. \tag{3}$$

We assert that there exist cycles $\Delta^n, \ldots, \Delta^{r+1}$ such that Δ^h is a generator of $\mathfrak{O}^h_p$ and

$$\rho_h \Delta^h = \delta \Delta^h = 0 \qquad (h = n, \ldots, r + 1).$$

Let Γ^n be a generator of the ordinary homology group $\mathfrak{H}^n\,(M, \mathfrak{Z}_p)$. Then $s\Gamma^n = k\Gamma^n$, k an integer. Evidently $k^p = 1 \bmod p$ so that $k = 1 \bmod p$. Hence $\delta\Gamma^n = 0$. Similarly $\rho_n\Gamma^n = 0$. This last relation implies that Γ^n is an element, evidently non-zero, of $\mathfrak{H}^n_p$. By (1) we have

$$\Gamma^n = \Delta^n + \Delta^n_0 (\Delta^n \subset M - L,\ \Delta^n_0 \subset L) \tag{4}$$

where Δ^n, Δ^n_0 are elements of $\mathfrak{O}^n_p$, $\mathfrak{H}^n_{0p}$, respectively. Since the latter group is zero, Δ^n cannot be the zero of $\mathfrak{O}^n_p$, hence is a generator. Since all chains in L, in particular Δ^n_0, are annulled[11] by δ and ρ_n, Δ^n is also annulled by δ, ρ_n and therefore has the required properties. Suppose that $\Delta^n, \ldots, \Delta^{h+1}$ $(h > r)$ have been chosen. The relations $\rho_{h+1}\Delta^{h+1} = \delta\Delta^{h+1} = 0$ imply (lemma 2) $\Delta^{h+1} = \rho_h \sigma Y^{h+1}$ say, where $Y \subset M - L$. Let $\Gamma^h = F(\sigma Y)$. Then

$$\rho_h \Gamma^h = F\rho_h \sigma Y = F\Delta^{h+1} = 0$$
$$\delta \Gamma^h = F\delta\sigma Y = 0.$$

Now Γ^h is precisely the image of Δ^{h+1} under a certain onto-homomorphism[12] $\alpha: \mathfrak{H}_p^{h+1} \to \mathfrak{H}_p^h$. Since these groups are both of order p, α is an isomorphism and so Γ^h is a generator of $\mathfrak{H}_p^h = \mathfrak{D}_p^h$. We then obtain the required Δ^h from Γ^h as for the case $h = n$.

We can now prove the theorem by obtaining the contradiction $\mathfrak{D}_p^{r+1} = 0$. As above with h replaced by r, we may write $\Delta^{r+1} = \rho_{r+1}\sigma Y^{r+1}$. Let $\Gamma^{r+1} = F\sigma Y^{r+1}$. As in (4) we have $\Gamma^r = \Delta^r + \Delta_0^r$ ($\Delta_0^r \subset L$) where Δ^r is an element of $\mathfrak{D}_p^r$. Since this group vanishes, there exists a special homology

$$F(\rho_r H^{r+1}) = \Delta^r.$$

The boundary of

$$\sigma Y^{r+1} - \rho_r H^{r+1} \tag{5}$$

is $\Delta_0^r \subset L$. This boundary, however, is of the form $\sigma A + \rho_r B$ hence its intersection with L is null.[13] Hence (5) is an absolute cycle and is therefore the boundary of a chain say H^{r+2}. Then

$$\begin{aligned} F\rho_{r+1}H^{r+2} &= \rho_{r+1}\sigma Y^{r+1} - \rho_{r+1}\rho_r H^{r+1} \\ &= \rho_{r+1}\sigma Y^{r+1} = \Delta^{r+1} \end{aligned}$$

so that Δ^{r+1} satisfies a special homology and therefore is the zero of $\mathfrak{D}_p^{r+1}$. Since Δ^{r+1} is a generator, we conclude that $\mathfrak{D}_p^{r+1} = 0$.

---

[1] Also called ρ-homologies (5 and 6). All references are to Appendix B of Lefschetz's *Algebraic Topology*.

[2] 1.

[3] The elements $t_1, \ldots, t_m$ are independent if $\Pi t_i^{a_i} = 1$ implies $a_i = 1$ modulo the order of t_i.

[4] 4.1.

[5] (2.1) Theorem.

[6] If $p > 2$, every periodic transformation of an even dimensional euclidean sphere admits fixed points. Hence the corollary is of interest only when n is odd.

[7] 9.

[8] 5 and 6.

[9] 7.1.

[10] 12.

[11] 4.3.

[12] 6 and 10.8.

[13] 4.3.

ON THE ARITHMETIC IN A GROUP RING

By Richard Brauer

University of Toronto[1]

Communicated March 24, 1944

1. Every group $\mathfrak{G}$ of finite order g determines an associative algebra Γ, the group ring, over an arbitrary field K. This algebra consists of all linear combinations

$$\alpha = \sum a_\iota G_\iota \tag{1}$$

of the group elements $\mathfrak{G}_\iota$ with coefficients a_ι in K, where equality of these elements, addition and multiplication are defined in the natural manner. The study of the algebra Γ forms one of the most powerful weapons we have for the investigation of groups of finite order. When the field K is of characteristic 0, the algebra Γ is semisimple, and the general principles of the theory of algebras can be applied. As is well known, the ordinary theory of group representations and group characters can be obtained in this way. We then assume that the field K is algebraically closed, or at least that K has the property that all irreducible representations of $\mathfrak{G}$ in K are absolutely irreducible. If the latter condition does not hold in the original field K, it will be satisfied in algebraic extension fields of K of finite degree. Without loss of generality, we may restrict ourselves to the case that K is a suitable algebraic number field,[2] as we shall assume throughout this note.

There can be little doubt that we are far from knowing all important properties of group characters. In particular, we are interested in further results which connect the group characters directly with properties of the abstract group $\mathfrak{G}$. Any result of this kind means, in the last analysis, a result concerning the structure of the general group of finite order.

2. One approach to our question is to study the arithmetic properties of Γ after we study the algebraic properties of Γ. An element α of Γ is said to be an integer, if all the coefficients a_ι in (1) are integers of K. The ring J of these integers is not a maximal order,[3] and therefore the ordinary theory of ideals does not hold. However, our definition of integer is linked with the group $\mathfrak{G}$ in a natural manner, and we would lose this connection, if we replaced J by a maximal order. Actually, a study of most arithmetic properties of a maximal order would not lead us beyond the algebraic properties of Γ.

The arithmetic question in which we are mainly interested[4] is the question how a prime ideal $\mathfrak{p}$ of K (or rather its extension $(\mathfrak{p}) = \mathfrak{p}J$ to J) behaves in J. This leads to the study of the residue class ring $J^* = J/(\mathfrak{p})$. But J^* can be interpreted as the group ring of $\mathfrak{G}$ over the field K* of residue classes of integers of K modulo $\mathfrak{p}$. We thus come back to a study of a group ring and of representations of $\mathfrak{G}$, but now with regard to a modular field K*. The investigation of $(\mathfrak{p})$ becomes equivalent to the study of the modular representations[5] of $\mathfrak{G}$. For every prime ideal divisor $\mathfrak{P}$ of $(\mathfrak{p})$, the residue class ring $J/\mathfrak{P}$ is a complete matric algebra of a certain degree f over K*. Thus, $\mathfrak{P}$ defines a modular irreducible representation $\mathfrak{F}$ of $\mathfrak{G}$ of degree f, and we have a (1-1)-correspondence between the prime ideal divisors $\mathfrak{P}_1, \mathfrak{P}_2, \ldots, \mathfrak{P}_k$ of $(\mathfrak{p})$, and the irreducible modular representations $\mathfrak{F}_1, \mathfrak{F}_2, \ldots, \mathfrak{F}_k$ of $\mathfrak{G}$.

Since J is not a maximal order, we cannot write $(\mathfrak{p})$ as a product of powers of the $\mathfrak{P}_i$. However, we still have a unique representation of $(\mathfrak{p})$ as a direct intersection of ideals

$$(\mathfrak{p}) = \mathfrak{M}_1 \cap \mathfrak{M}_2 \cap \ldots \cap \mathfrak{M}_t. \qquad (2)$$

Here, any two of the $\mathfrak{M}_\tau$ are relatively prime, and no $\mathfrak{M}_\tau$ can be written as an intersection of relatively prime ideals which are different from $\mathfrak{M}_\tau$. The $\mathfrak{P}_i$ dividing a fixed $\mathfrak{M}_\tau$ form a "block" B_τ, and according to what we said above, we may also speak of the modular representations $\mathfrak{F}_i$ belonging to the block B_τ. Every $\mathfrak{F}_i$ then belongs to exactly one block. Finally, we may associate every ordinary irreducible representation $\mathfrak{Z}_j$ of $\mathfrak{G}$ with exactly one block B_τ.[6] Every $\mathfrak{Z}_j$ of the block B_τ contains only modular constituents $\mathfrak{F}_i$ of B_τ, and, conversely, every $\mathfrak{F}_i$ of B_τ appears only in $\mathfrak{Z}_j$ of B_τ. This shows that there exists a connection between the decomposition (2) and the decomposition of Γ into a direct sum of simple algebras.

Denote by p the rational prime number which is divisible by $\mathfrak{p}$. If p^α is the highest power of p which divides all degrees f_i of the $\mathfrak{F}_i$ belonging to the block B_τ, then p^α also is the highest power of p which divides the degrees z_j of all ordinary representations $\mathfrak{Z}_j$ of B_τ.[7] We then say that the block B_τ is of *type* α. If p^a is the highest power of p dividing the order g of $\mathfrak{G}$, then we call $d = a - \alpha$ the *defect* of B_τ. The smaller d, the simpler is the structure of $\mathfrak{M}_\tau$ in (2).

If the group characters of $\mathfrak{G}$ are known, we can easily obtain the blocks $B_1, B_2, \ldots, B_t$. The main object of this paper is to give a direct characterization of the blocks. For instance, we shall show that the number of blocks of given positive defect d is completely determined by the structure of certain subgroups $\mathfrak{N}$ of $\mathfrak{G}$ which are the normalizers $\mathfrak{N}$ of the subgroups $\mathfrak{H}$ of order p^d in $\mathfrak{G}$.

3. The modular group ring $J/(\mathfrak{p})$ determines a matrix C of Cartan invariants.[8] This C is the direct sum of the matrices C_τ of Cartan invariants of the rings $J/\mathfrak{M}_\tau$ corresponding to the blocks B_τ. The degree of C_τ is equal to the number y_τ of prime ideals $\mathfrak{P}_i$ (and of modular representations $\mathfrak{F}_i$) in the block. The matrix C_τ is symmetric, the corresponding

quadratic form is non-negative, the determinant of C_τ is a power of p. The coefficients c_{ij} are non-negative rational integers, the value of c_{ij} depends on the mutual relation of the prime ideals $\mathfrak{P}_i$ and $\mathfrak{P}_j$.

It is easy to determine the elementary divisors of C. Let $K_1, K_2, \ldots, K_k$ be the classes of conjugate elements of $\mathfrak{G}$ which contain *p-regular elements*, i.e., elements whose order is prime to p. Let g_λ be the number of elements in K_λ so that $g_\lambda = g/n_\lambda$ where n_λ is the order of the normalizer of the elements of K_λ, and suppose that p divides n_λ exactly with the exponent ρ_λ. Then $p^{\rho_1}, p^{\rho_2}, \ldots, p^{\rho_k}$ are the elementary divisors of C.[9] Exactly y_τ of these elementary divisors must belong to C_τ, and naturally the question arises in what manner the k elementary divisors of C are distributed into blocks. A partial answer is given by

THEOREM 1: *If C_τ is the Cartan matrix of a block of defect d, then C_τ has one elementary divisor p^d while all other elementary divisors of C_τ are powers of p with exponents smaller than d.*

As corollaries[10] we obtain

COROLLARY 1: *If there exist l_σ p-regular classes K_λ in $\mathfrak{G}$ for which the order of the normalizer of the elements is divisible by p^σ but not by $p^{\sigma+1}$, then $\mathfrak{G}$ possesses at most l_σ blocks of defect σ ($\sigma = 0, 1, 2, \ldots, a$).*

COROLLARY 2: *There exist exactly l_a blocks of maximal defect a, where $g \equiv 0 \pmod{p^a}$, $g \not\equiv 0 \pmod{p^{a+1}}$.*

COROLLARY 3: *If B_τ is a block of defect 0, then C_τ has the degree 1, and its only coefficient is 1. The block consists of exactly one ordinary representation $\mathfrak{Z}$ and one modular representation $\mathfrak{F}$. Taken as a modular representation, $\mathfrak{Z}$ remains irreducible and coincides with $\mathfrak{F}$. The prime ideal $\mathfrak{P}$ belonging to $\mathfrak{F}$ coincides with $\mathfrak{M}_\tau$ in (2).*

In the proof of theorem 1, the following construction can be used. Consider the k columns of the matrix of group characters of $\mathfrak{G}$ which belong to the p-regular classes $K_1, K_2, \ldots, K_k$ of $\mathfrak{G}$. It is possible to select a minor of degree k which is not divisible by $\mathfrak{p}$. Exactly y_τ of the rows in this minor must belong to ordinary characters ζ_μ of B_τ. It is then possible to associate y_τ of the columns with B_τ, such that each column belongs to exactly one B_τ, and that the y_τ rows and columns associated with B_τ form a minor which is not divisible by $\mathfrak{p}$. (This construction may not be unique.) We can then prove

THEOREM 2: *If the classes $K_\lambda, K_\mu, K_\nu, \ldots$ are associated with the block B_τ by the preceding construction, and if p divides the order of the normalizers of the elements of K_i to the exact exponent ρ_i, then the elementary divisors of C_τ are the powers of p with the exponents $\rho_\lambda, \rho_\mu, \rho_\nu, \ldots$.*

If ζ_μ is an ordinary character of $\mathfrak{G}$, and if $\zeta_\mu(K_\nu)$ is the value of ζ_μ for the elements of K_ν, then it is well known that the numbers

$$\omega_\mu(K_\nu) = g_\nu \zeta_\mu(K_\nu)/Dg(\zeta_\mu) \tag{3}$$

are algebraic integers. The proof of theorems 1 and 2 yields at the same time

THEOREM 3: *A character* ζ_μ *belongs to a block* B_τ *of a defect larger than or equal to a given number d, if and only if we have* $\omega_\mu(K_\nu) \equiv 0 \pmod{\mathfrak{p}}$ *for all p-regular classes with* $\rho_\nu < d$. *Two characters* ζ_λ *and* ζ_μ *belonging to blocks of defect d appear in the same block, if and only if*

$$\omega_\lambda(K_\nu) \equiv \omega_\mu(K_\nu) \pmod{\mathfrak{p}}$$

for all p-regular classes with $\rho_\nu = d$.

COROLLARY 4: *If the Cartan matrix* C_τ *of the block* B_τ *of defect d has h elementary divisors* p^{d-1}, *then* B_τ *contains at least* $h + 1$ *ordinary characters whose degrees are not divisible by* p^{a-d+1}. (By definition, the degrees are all divisible by p^{a-d}.)

4. If K_α is a class of conjugate elements of $\mathfrak{G}$, we also denote by K_α the sum of all the elements in this class. As is well known, the K_α then form a basis of the center Λ of the group ring Γ. We thus have formulae

$$K_\alpha K_\beta = \sum a_{\alpha\beta\gamma} K_\gamma \tag{4}$$

where the $a_{\alpha\beta\gamma}$ are non-negative rational integers. The true significance of the numbers ω_μ in (3) is that they form a character of the commutative algebra Λ. Taking the $\omega_\mu \pmod{\mathfrak{p}}$, we obtain a modular character of Λ, or what is the same thing, a character of the center Λ^* of the modular group ring Γ^*. But characters ζ_μ of $\mathfrak{G}$ belonging to the same block B_τ give rise to the same modular character of Λ. Hence we have a (1-1) correspondence between the blocks B_τ of $\mathfrak{G}$ and the characters of the commutative modular algebra Λ^*.

It is easy to prove certain arithmetic properties of the $a_{\alpha\beta\gamma}$ in (4). For instance, if an element G of K_γ commutes with all elements of a subgroup $\mathfrak{H}$ of order p^d of $\mathfrak{G}$ while no element of K_α commutes with all elements of $\mathfrak{H}$, then $a_{\alpha\beta\gamma}$ must be divisible by p. On the basis of this remark and more general statements of a similar nature, it is possible to establish connections between $\Lambda^* = \Lambda^*(\mathfrak{G})$ and the analogous algebras $\Lambda^*(\mathfrak{N})$, formed by means of suitable subgroups $\mathfrak{N}$ of $\mathfrak{G}$. In this manner we obtain a connection between the characters of $\Lambda^*(\mathfrak{G})$ and those of $\Lambda^*(\mathfrak{N})$, and finally a connection between the blocks of $\mathfrak{G}$ and certain blocks of $\mathfrak{N}$. Our main results are proved by combining this method with that of section 3. They are contained in the following theorems:

THEOREM 4: *Choose one group* $\mathfrak{H}_\sigma$ *of order* p^d *from each class of conjugate subgroups of this order of* $\mathfrak{G}$, *and denote the normalizer of* $\mathfrak{H}_\sigma$ *by* $\mathfrak{N}_\sigma$ $(\sigma = 1, 2, \ldots, l)$. *For the given prime p, all blocks of* $\mathfrak{N}_\sigma$ *have at least the defect d. If* $\mathfrak{N}_\sigma$ *possesses exactly* r_σ *blocks of defect d, then* $\mathfrak{G}$ *contains exactly* $r_1 + r_2 + \ldots + r_l$ *blocks of defect d.*

THEOREM 5: *To every block B of defect d of $\mathfrak{G}$ there corresponds uniquely a block $\tilde{B}$ of defect d of one of the $\mathfrak{N}_\sigma$ in theorem 4. If $\omega_\mu(K_\nu)$ in (3) is formed by means of a character of B while $\tilde{\omega}_\lambda(\tilde{K}_\iota)$ is formed in an analogous manner by means of a character of $\tilde{B}$, then*

$$\omega_\mu(K_\alpha) \equiv \sum_\beta \tilde{\omega}_\nu(\tilde{K}_\beta) \ (mod\ \mathfrak{p}). \tag{5}$$

Here, K_α is any class of $\mathfrak{G}$, while $\tilde{K}_\beta$ ranges over all classes of $\mathfrak{N}_\sigma$ which lie in K_α and whose elements belong to the centralizer $\mathfrak{T}_\sigma$ of $\mathfrak{H}_\sigma$. The ideal $\mathfrak{p}$ denotes a fixed prime ideal divisor of p in the field generated by the characters of B and $\tilde{B}$.

COROLLARY 5: *If in the notation of theorem 5, the class K_α does not contain elements of $\mathfrak{T}_\sigma$ then*

$$\omega_\mu(K_\alpha) \equiv 0 \ (\text{mod } \mathfrak{p}) \tag{6}$$

If K_α contains elements of $\mathfrak{T}_\sigma$, and if the order of the normalizer of these elements is not divisible by p^{d+1}, then K_α contains only one class $\tilde{K}_\beta$ which consists of elements of $\mathfrak{T}_\sigma$, and we have

$$\omega_\mu(K_\alpha) \equiv \tilde{\omega}_\nu(\tilde{K}_\beta) \ (\text{mod } \mathfrak{p}). \tag{7}$$

The congruences (5), (6) and (7) can be expressed in terms of the characters. We thus obtain the following relations between the characters of $\mathfrak{G}$ and those of $\mathfrak{N}_\sigma$.

COROLLARY 6: *Let ζ_μ by any character of B, say of degree z, let $\tilde{\zeta}_\nu$ be a character of the corresponding block $\tilde{B}$ of $\mathfrak{N}_\sigma$, say of degree w. In the notation of theorem 5, we have*

$$\zeta_\mu(K_\alpha) \equiv \frac{z}{wg_\alpha} \sum_\beta \tilde{g}_\beta \tilde{\zeta}_\nu(\tilde{K}_\beta) \ (\text{mod } \mathfrak{p}p^m) \quad \text{with } m = s - a + \rho_\alpha$$

where g_α is the number of elements in K_α, $\tilde{g}_\beta$ is the number of elements in $\tilde{K}_\beta$, ρ_α is the exponent to which p divides $n_\alpha = g/g_\alpha$, and s is the exponent to which p divides z. In the case of (6), this yields

$$\zeta_\mu(K_\alpha) = 0 \ (\text{mod } \mathfrak{p}p^m),$$

while in the case of (7) we obtain

$$\zeta_\mu(K_\alpha) \equiv \frac{zn}{wg} \tilde{\zeta}_\nu(\tilde{K}_\beta) \ (\text{mod } \mathfrak{p}p^{s-a+d})$$

where n is the order of $\mathfrak{N}_\sigma$.

Our results do not give any new information about the blocks of highest kind, i.e., of defect 0. We have here $\mathfrak{H} = \{1\}$, $\mathfrak{T}_\sigma = \mathfrak{N}_\sigma = \mathfrak{G}$. For blocks of positive defect d, we may have $\mathfrak{N}_\sigma = \mathfrak{G}$, but only when $\mathfrak{G}$ has a normal subgroup of an order $p^d > 1$. In any case, every block of defect d of $\mathfrak{N}_\sigma$

contains an ordinary representation which associates the unit matrix with the elements of $\mathfrak{H}_\sigma$. As a representation of $\mathfrak{N}_\sigma/\mathfrak{H}_\sigma$ this one is of the highest kind. This shows that the question of finding all blocks of defect $d > 0$ can be answered by studying groups of smaller order than g. If for every one of these B we know the number of ordinary characters contained in B, we also know the number of blocks of defect 0. By corollary 3, it is the number of classes of $\mathfrak{G}$, diminished by the number of ordinary characters which belong to blocks of positive defect.

In finding all blocks of $\mathfrak{G}$, the following remark is useful.

REMARK: In theorems 4 and 5, it is sufficient to consider only groups $\mathfrak{H}_\sigma$ which are p-Sylow-subgroups of normalizers of p-regular elements, and for which $\mathfrak{H}_\sigma$ is a maximal normal p-subgroup of $\mathfrak{N}_\sigma$. If these assumptions are not satisfied, $\mathfrak{H}_\sigma$ does not possess blocks of defect d.

---

[1] The work on this note was done while the author was a Fellow of the John Simon Guggenheim Memorial Foundation.

[2] Using a theorem of Hasse, *Jour. f. d. reine u. angew. Math.*, **167**, 399–404 (1931), we could even assume that K was obtained from the field of rational numbers by the adjunction of a certain root of unity.

[3] For the arithmetic in maximal orders of algebras, see, for instance, Jacobson, N., "The Theory of Rings," Math. Surveys No. II (1943).

[4] For the following cf. Brauer, R., these PROCEEDINGS, **25**, 252–258 (1939).

[5] For the modular representations of groups, cf. Brauer, R., and Nesbitt, C., University of Toronto Studies, Math. Ser. No. 4 (1937), and *Ann. Math.*, **42**, 556–590 (1941). I refer to the second of these papers as BN.

[6] Cf. BN, §9.

[7] BN, §17.

[8] BN, § 5.

[9] BN, § 16.

[10] Corollary 1 is an improvement of theorem 3 of BN. Corollary 2 is identical with theorem 2 of BN. Corollary 3 is identical with theorem 1 of BN.

POSSIBLE NUMBERS OF NON-INVARIANT OPERATORS OF A GROUP

BY G. A. MILLER

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ILLINOIS

Communicated April 8, 1944

In a preceding number of these *Proceedings* (**30**, 25–28 (1944)), it was noted that a necessary and sufficient condition that there is a group in which the number of the non-invariant operators is a given prime number p is that there is a group of order $p + 1$ which contains no invariant operator besides the identity. From this theorem it follows that the only odd prime numbers which are less than 100 and are not equal to the number of the non-invariant operators in some group are 3, 7, 31, 43. It is obvious that there is no group which contains exactly two or four non-invariant operators and if a group contains exactly five non-invariant operators it is the non-cyclic group of order 6. Each of the two non-abelian groups of order 8 contains exactly six non-invariant operators and these are the only groups which have this property.

From the noted theorem it results that whenever a prime number is of the form $2^n - 1$ it cannot represent the number of the non-invariant operators of any group. In particular, there is no group which contains exactly seven non-invariant operators. There is also no group which contains exactly eight non-invariant operators since the smallest number of non-invariant operators in a non-abelian group of order g is known to be $3g/4$ and the non-abelian group of order 10 contains nine non-invariant operators. The tetrahedral group is known to contain eleven non-invariant operators while each of the other two non-abelian groups of order 12 clearly contains exactly ten non-invariant operators. We have now noted all the possible non-abelian groups which separately contain no more than eleven non-invariant operators.

Suppose that the order of the central quotient group of a given group G is a power of a prime number. A Sylow subgroup of G whose order is a power of this prime number contains an invariant subgroup whose order is this prime number. The operators of this invariant subgroup are commutative with every operator in the central of G and also with at least one operator of G corresponding to every operator in the central quotient group G. Hence these operators appear in the central of G, which is contrary to the hypothesis unless the central is divisible by the prime number which divides the order of this central quotient group. Hence we have proved the following theorem: *When the order of the central quotient group of G is a certain power of a prime number then the order of G is divisible by a higher power of this prime number and hence the order of the central of G is divisible by this prime number.*

If a given group G contains exactly n non-invariant operators then the direct product of G and an arbitrary abelian group contains exactly n times the order of this abelian group invariant operators. In particular, from the fact that there is a group which contains exactly five non-invariant operators it results directly that there is at least one group which contains an arbitrary multiple of five non-invariant operators. We have seen that there are two and only two groups which separately contain ten non-invariant operators but it is obvious that there is only one group which contains exactly fifteen non-invariant operators since *a group whose central involves one of its Sylow subgroups and whose order is divisible by only two distinct prime numbers is the direct product of its Sylow subgroups.* In particular, the central of a group of order 20 cannot be of order 5.

It was noted above that both the dihedral group of order 6 and the dihedral group of order 10 are completely determined by the number of their respective non-invariant operators. To construct a system of groups which has the property that every group contained in the system is completely determined by the number of the non-invariant operators contained in the group we may note that whenever $2p - 1$ is a prime number, where p is also a prime number, the dihedral group of order $2p$ contains a prime number of non-invariant operators. Since there is one and only one dihedral group of order $2p$ and this contains no invariant operator besides the identity when p is odd it results that each of these groups is completely determined by the number of its non-invariant operators. The three lowest orders of such groups are 6, 14, 38.

Although there are many groups which are completely determined by the number of their non-invariant operators and this number therefore serves as a complete definition of these groups there are also many cases in which different groups have the same number of non-invariant operators. It was noted above that the two non-abelian groups of order 8 and two of the three non-abelian groups of order 12 have this property. Among the nine non-abelian groups of order 16 there are six which have a central of order 4 and hence they contain exactly twelve non-invariant operators. The three remaining non-abelian groups of order 16 contain a central of order 2 and hence involve separately fourteen non-invariant operators in each case. Since the groups of order 16 have been widely published it may be assumed that these cases are known. They represent the only groups which contain exactly either twelve or fourteen non-invariant operators.

We have now incidentally noted that each of the natural numbers which are less than 16 except 1, 2, 3, 4, 7, 8 is equal to the number of the non-invariant operators in at least one group. To prove that there is no group which has exactly sixteen non-invariant operators we may first note that the order of such a group could not be 18 for it would then contain a central of order 2 and hence it would contain only one subgroup of this order and this would be a Sylow subgroup. For similar reasons it could not be of order 20. It could not be of a larger order since at least three-fourths of the operators of every non-abelian group are non-invariant. There exists therefore no group which contains exactly sixteen non-invariant operators. In fact, there exists no group which has exactly a power of 2 less than 2^7 non-invariant operators.

If a group contains exactly seventeen non-invariant operators it is either the dihedral or the generalized dihedral group of order 18 and if it contains exactly eighteen non-invariant operators its order could not exceed 24. If it is of order 24 its central must contain its Sylow subgroup of order 3 and hence it must be the direct product of a group of order 3 and one of the two non-abelian groups of order 8 according to the theorem noted above.

If it is of order 20 it is either the dihedral or the dicyclic group of this order and hence there are exactly four groups which have the property that each of them contains exactly eighteen non-invariant operators. Since a group which involves exactly nineteen non-invariant operators is the metacyclic group of order 20 it is completely determined by the number of its non-invariant operators and the two noted groups of order 18 are the only groups up to order 20 which contain exactly a prime number of non-invariant operators but are not completely determined by this number.

The alternating group G of order 60 is known to be the group of smallest composite order which is simple. We proceed to prove that this group is completely determined by the fact that it contains exactly fifty-nine non-invariant operators. To prove this fact it may first be noted that since fifty-nine is a prime number the number of the groups which contain exactly fifty-nine non-invariant operators is equal to the number of groups of order 60 which separately have the property that each of them contains no invariant operator besides the identity. If a group of order 60 contains no invariant operator besides the identity it involves six subgroups of order 5 which it transforms according to a transitive group of degree 6. This transitive permutation group must be of order 60 since each of its six subgroups of order 5 transforms the remaining five subgroups of this order contained in G among themselves but not invariantly.

Hence, the permutation group according to which G transforms its six subgroups of order 5 is simply isomorphic with G and there is therefore only one permutation group of degree six and of order 60 which is simply isomorphic with G, and hence the simple group of order 60 is completely determined by the fact that it contains fifty-nine non-invariant operators. The fact that the alternating group of degree five is the simple group of smallest composite order was noted already by E. Galois (1811–1832). The fact that this group is completely determined by the number of its non-invariant operators furnishes a new definition of this important simple group of composite order. Various other definitions of this group are known.

THE BIOLOGICAL EFFECTS OF X-RAYS AS A FUNCTION OF INTENSITY

By A. A. Bless

Department of Physics, University of Florida

Communicated April 3, 1944

The factors which are influencing the biological action of radiations naturally received a great deal of attention. The prevailing opinion seems to be that the biological effects of radiations are a function only of the energy of the absorbed rays. This view seems to be reasonably well established for radiations of x-ray frequency.[1] In the range of frequencies of the band of ultra-violet radiations the effect seems to depend on the particular wave length.[2]

The effects of other radiations, such as neutrons or alpha particles appears to be also a function of the total energy absorbed, though in this case the energy of the individual particles seems to be a factor as well.[3] However, the effects of a given amount of energy of neutron rays are different from the effects produced by the same amount of absorbed x-ray energy. The ratio of effectiveness of a given amount of neutron ray energy is many times as great as the absorption of the same amount of x-ray energy. This ratio does not at all remain constant but varies for different effects and with different specimens. In the case of mutation of drosophila the ratio is 2:1; for the effects on carcinoma, is 6:1; for the recessive effects on wheat the ratio is 20:1 and so on.[4]

The difference in the action of the two types of radiation is, in the opinion of Zircle[3] and of Failla,[1] due to the greater intensity of radiation produced by neutron and alpha rays compared with the intensity of ionization produced by x-rays. Zircle points out that the ratio of the intensities of the two types of radiation is of the order of 100:1. In very careful experiments Zircle showed that the biological effects of alpha particles vary approximately with the 2.5 power of that intensity.

The view that the intensity of ionization is responsible for the difference in the biological action of the various radiations seems to have a great deal in its favor. The seat of the action of x-rays or of other types of radiations are the cells or some parts of the cell. The exchange of nutrients of this part of the cell with the rest of the organism must depend upon the boundary of the affected region. The ability to exchange products with the neighboring regions is profoundly affected by the electrical conditions in that boundary. It seems, therefore, reasonable to assume that high ionization near the boundary of a given region may profoundly affect the biological action of that region.

In the case of x-rays no intensity effects have been observed so far as the author is aware. It is possible that the x-ray energy was in all cases administered at such low intensities that the effects escaped the attention of the workers.

It seemed worth while to investigate whether the biological effects of a given dose of x-rays administered at a very high intensity are different from the effects produced by an equal dose at moderate intensities.

The source of x-rays was a deep therapy tungsten target water cooled Coolidge x-ray tube with a 45 degree focus. The tube was run at about 70,000 volts and 10 milliamperes. For most measurements the radiation was filtered with 1.2 mms. of aluminum. All these factors were kept constant in order to have the same band of radiations. The only factor that was varied was the distance from the tube. The construction of the tube was such that specimens could be placed at a distance of 6 cms. from the target. The distance was varied from 60 cms. to 6 cms. with a variation of the intensity of radiation from 25 to 2500 r/min. No experiments were made at longer distances from the target.

In order to increase the range of intensities, experiments were also conducted with unfiltered radiation at the smallest distance from the target, namely 6 cms. Since the quality of the unfiltered radiation was different from the filtered it seemed advisable to determine the relative effectiveness of the two types at low intensities in order that the intensity effect, if present, should not mask the comparison. The effect on seeds of wheat of unfiltered and filtered radiation were observed at a distance of 60 cms. from the target. The radiation at this distance was, doubtless, too weak to have an intensity effect. It was found that the time necessary to produce a given effect (per cent growth) with unfiltered rays was approximately one-half that of the filtered. The intensity of the unfiltered radiation at a distance of 6 cms. from the target was thus about 5000 r/min. The experiments were run for eight different intensities of radiation ranging from 25 r/min. to 5000 r/min.

The choice of a suitable biological material proved to be most difficult. Many unsuccessful experiments have been made with beans, peas and corn. Unfortunately the variation between individual species was much too great. Also, the variation of the effects of a given dose of x-rays even when the seeds were inbred was so great as to render the results doubtful, except for a prohibitively large number of measurements. The seeds of wheat proved to be more free from these defects than any other material tried and were finally adopted for the purpose. The Big Club variety was selected chiefly because some work has already been done on these seeds.

The experiments extended for a period of over two years. While the results obtained during the first year were quite consistent, the wheat seedlings were very defective that year, and it seemed advisable to repeat the experiment with a new and more vigorous crop. This was done last fall. The procedure followed was that employed by Zircle and Aebersold.[4]

The seeds were allowed to soak in water for 3 hours. They were then spread in trays over moist filter paper. After about 12 hours the seeds were placed on small plates, 30 seeds on each plate, and exposed to x-rays. After 48 hours of sprouting, the length of the stems was measured, and this length was used as the biological indicator. During Winter the variation in the temperature of the rooms was too great to follow this procedure. In some cases the rooms were so cold that even after 48 hours the growth was very small; on the other hand the growth was excessive at other times when the rooms were over-heated. Instead of following the practice of measuring the length of the seeds 48 hours after exposure to x-rays, the procedure was adopted of examining the seeds only after the average of the controls reached a height of about 45 mms. Later experiments were conducted with the seeds in a constant temperature enclosure. This method gave the best results so far as the magnitude of the spread from the mean is concerned.

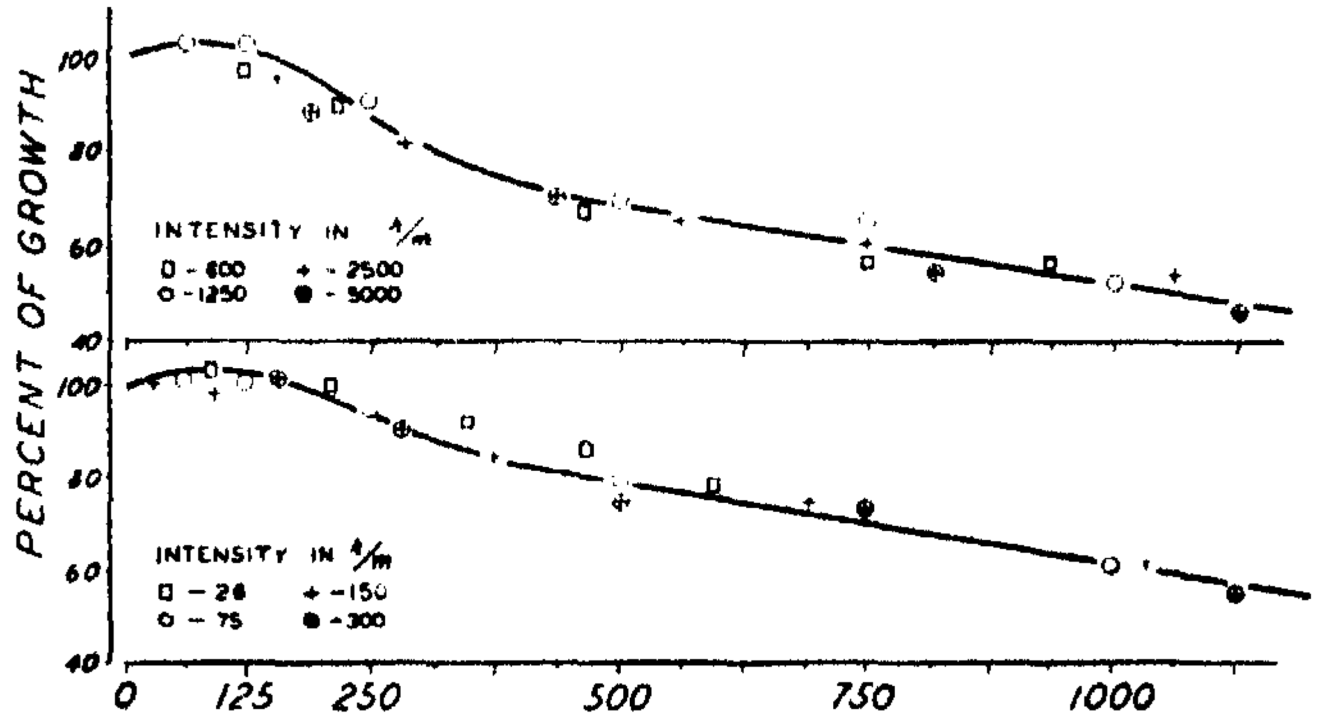


FIGURE 1

The variation of the length of the stem was used as the criterion of the effects of radiation, and no attention was paid to the roots. The reason for neglecting these data was the great variability of the root system. In some plants the roots were short and stubby, while in others they were long and slender. Also the number differed with different plants—some had two, while others had three or four.

The variation of the length of the stem as a function of the dose was studied at various intensities of x-rays. Data were also taken for a given dose at intensities varying from the smallest to the largest. Figure 1 shows the variation of the per cent of growth as a function of the dose with radiations at various intensities. It was subdivided into two curves in order to show more conveniently the results of the measurements with the eight different intensities. Each point represents the average of about 200 measurements. It is evident that the various intensities do not produce any marked differences in the results.

In figure 2 are shown the variations of the per cent growth for a given dose with the intensity of radiation. In these curves are incorporated data obtained in figure 1 and also data of experiments which were run especially for this purpose. Again it is evident that no significant variations with intensity are present.

It is evident that within the limits employed here the effects of x-rays do not depend at all on the intensities at which a given dose is administered. It is possible that the density of ionization produced by x-rays even at 5000 r/min. is still much too small compared to density of ionization produced by alpha particles or by protons. It is also possible that the character of the action of the two types of radiation is different.

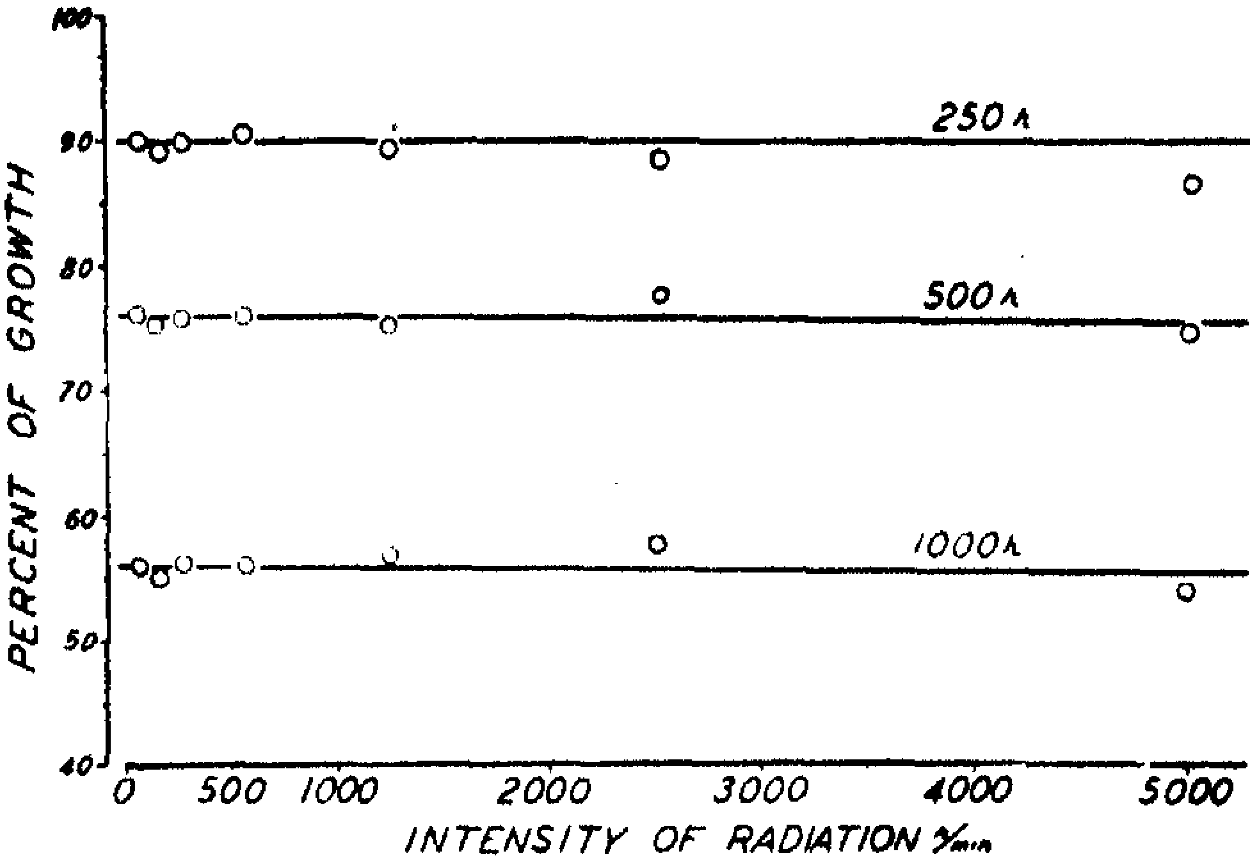


FIGURE 2

---

[1] G. Failla, *Jour. Appl. Phys.*, **12**, 279 (1941) Arthur A. Bless, *Proc. Nat. Acad. Sci.* **23**, 194 (1937).

[2] Florence E. Meier, *Smiths. Misc. Coll.*, **95**, 2 (1936).

[3] R. E. Zircle, *Am. Jour. Canc.*, **23**, 558 (1933).

[4] R. E. Zircle and P. C. Aebersold, *Proc. Nat. Acad. Sci.*, **22**, 134 (1936). John H. Lawrence and Robert Tenant, *Jour. Exp. Med.*, **66**, 667 (1937). R. E. Zircle and Isadore Lampe, *Am. Jour. of Roent. and Rad. Ther.*, **39**, 613 (1938). *A. Marshak Proc. Soc. Exp. Biol. and Med.*, **41**, 176 (1939).

PROCEEDINGS
OF THE
NATIONAL ACADEMY OF SCIENCES

Volume 30 June 15, 1944 Number 6

Copyright 1944 by the National Academy of Sciences

THE EFFECT OF X-RAYS UPON DOMINANT MUTATION IN MAIZE[1]

BY L. J. STADLER

DIVISION OF CEREAL CROPS AND DISEASES, U. S. DEPARTMENT OF AGRICULTURE, AND MISSOURI AGRICULTURAL EXPERIMENT STATION

Communicated April 13, 1944

No dominant mutations have been found in x-ray progenies of barley and maize, in the course of experiments in which many hundreds of recessive mutations have been detected. Many dominant alterations appear in the first generation progeny produced by the use of x-rayed pollen, but these alterations are lethal to the gametophyte and are therefore not reproduced in later generations. The recessive mutations detected are at least partially haplo-viable, since they must be transmitted through both male and female gametophytes before their effects may be observed in homozygous individuals. Since all of the induced alterations which are viable in haplophase are wholly recessive, these experiments provide no evidence of the production by x-rays of any mutant gene of positive action. The induced mutations observed may be merely gene losses tolerated by the gametophyte.[2,3]

The evidence against the occurrence of x-ray-induced dominant mutation is, however, inconclusive, for the following reasons:

1. The number of genes capable of showing the effects of dominant mutation may be much smaller than the number capable of showing recessive mutation, since many genes may be fixed by natural selection at a level maximal for phenotypic effect. The preponderance of recessive mutations may be only a reflection of the preponderance of such alleles in the genotype of the irradiated individual. The possibility of inducing dominant mutation therefore may be tested critically only by the effects upon known recessives, that is, genes with known dominant alleles.

2. Among known recessives many may be themselves deficiencies and therefore incapable of mutation. Critical evidence of failure to mutate to a dominant allele therefore may be obtained only from recessive genes which have previously been known to mutate to dominant alleles.

3. The only recessive alleles which meet this requirement are the so-

called variegation genes, which may be regarded as unstable recessives mutating frequently to a dominant allele. Each of these is characterized by a high spontaneous frequency of dominant mutation, so high that the effect of x-ray treatment in inducing additional dominant mutations probably would be inappreciable. Demerec's experiment[4] with the unstable miniature-3 of *Drosophila virilis* is an example of this difficulty.

It is possible to avoid these difficulties in the case of one gene. The recessive gene a_1 in maize has several known dominant alleles. Rhoades[5,6] has shown that in the presence of another gene, *Dt*, mutations of *a* to various dominant alleles occur. With increasing dosage of *Dt* the frequency of dominant mutation of *a* is sharply increased, and in endosperms triplex for *Dt* and for *a* (*a a a Dt Dt Dt*) more than 100 mutations per seed are usually found. In the absence of *Dt* no dominant mutations of *a* have been reported, though *a a a dt dt dt* endosperms have been closely examined in large numbers. (Dr. Rhoades informs me that he has seen one instance of an apparent "dot" of anthocyanin colored tissue in the aleurone cells of an *a a a dt dt dt* seed.)

The gene *a*, then, is capable of mutation to a recognizable dominant allele, and can be tested for response to x-ray treatment in genotypes in which spontaneous dominant mutations of the gene rarely if ever occur. Since the effects of mutation are clearly recognizable in minute sectors of tissue, the treatment may be applied at a fairly advanced stage in endosperm development, so that many hundreds of cells may be tested for mutation by examination of a single endosperm. It is possible therefore to test for the occurrence of this mutation in practically unlimited populations.

The experiment was conducted as follows: The seed to be irradiated was produced by the cross *a a* × *A a*, both parent stocks being homozygous for *dt dt* and for the complementary factors required for aleurone color development. The endosperms of half of the seeds produced are *A a a* (colored aleurone). These serve to indicate the size of the sectors resulting from genetic alterations induced by irradiation at the stage of development chosen, since induced deficiencies or recessive mutations of *A* in these endosperms result in sectors deficient in aleurone color. In the colorless seeds (*a a a*) induced dominant mutation of any one of the *a* genes would result in a corresponding sector of colored aleurone. The colored seeds thus provide a basis for estimating the number of opportunities for the occurrence of detectable mutation in the colorless seeds, and for comparison of the relative frequency of induced dominant mutation and induced loss of *A* due to recessive mutation or deficiency.

X-ray treatment was applied to the ears 73–81 hours after pollination, following preliminary experiments which indicated that treatment at this stage results in sectors of suitable size. Treatments were made on field-grown plants by means of a mobile x-ray unit with self-rectifying tube, operated at 88 K.V.P., 4.5 ma. tube current, and 15 inches target distance, with no filtration other than that of the overlying tissue. Under these conditions the time of treatment is 2.48 minutes per 100 r. The doses applied, uncorrected for filtration, were 800 r and 1600 r.

The lower dose was applied to 21 ears, which yielded a total of 8003 seeds. The effect of the treatment was clearly greater on the exposed side of the ear than on the opposite side, as indicated by the frequency of deficiency sectors and also by the occurrence of "scarred" endosperms. There was also appreciable variation in dosage along the long axis.

The average size and frequency of sectors showing "*A*-loss" (that is, loss of the *A* phenotype), was estimated by detailed examination of a small sample of seeds. From each of six representative ears given the lower dose of radiation, three seeds were taken from the treated side and three from the opposite side, with approximately even distribution along the long axis of the ear. Each of the 36 seeds was closely examined under magnification by the use of a dissecting microscope. This permits determination of the size of sectors by direct count of aleurone cells affected. In small sectors due to loss of *A* the aleurone cells are not wholly colorless, but the contrast in color is sharp enough to make the outline of the sector quite definite. In large sectors, produced by irradiation of the ear about 30 hours after pollination, the aleurone cells of sectors resulting from loss of *A* are colorless except for a rim of faintly colored cells about the margin of the sector. There are in addition occasional sectors of dilute aleurone color, probably resulting from similar changes affecting an unknown gene or genes. These are readily distinguishable from sectors involving *A* in the early-treated material, but are less distinct when the sectors are so small that marginal color may cover the entire sector. In order to minimize the error from this source, all sectors which seemed possibly due to loss of genes other than *A* were omitted. These made up nearly one-third of the aleurone sectors observed on the treated seeds. The ratio of dilute sectors to colorless sectors on early-treated seeds of the same cross is considerably less than 1:2, and it is therefore probable that the actual frequency of *A*-loss in the experimental material is somewhat higher than the data below would indicate. The observations are summarized in table 1.

The 18 seeds representing the exposed side of the treated ears (seeds Nos. 1 to 3 on each ear) yielded 618 sectors, or an average of 34 per seed; while those representing the opposite side yielded 248, or an average of 14 per seed. The sectors varied rather widely in size, the average number of aleurone cells included being 112. The frequency of *A*-loss sectors on untreated ears, determined by similar examination of 10 representative seeds from each of the 19 control ears, was 0.12 per seed. The 23 sectors found ranged in size from 9 to 750 cells.

The approximate number of aleurone cells in the entire endosperm was determined similarly by counts of aleurone cells in representative areas. The number of aleurone cells in seeds of average size was found to be about 160,000. It may be roughly estimated, therefore, that at the time of treatment the number of cells per seed whose genetic alteration could be detected by sectors in the mature aleurone is about 1400 (160,000/112).

The colorless seeds of the same treated ears permit the detection of mutations of *a* to any one of the various *A* alleles producing aleurone color. Each endosperm cell contains three *a*'s, each of which is capable, under the influence of *Dt*, of mutating to a colored aleurone allele. If the phenotypic effects of mutation and of deficiency are shown by the total cell progeny of the irradiated cell, or by the same fraction of this progeny, the result of mutation should be a colored sector equal in area to the color-deficient sectors produced by deficiency.

If either alteration requires for its realization a period of one or more cell generations in excess of the other, the resulting sectors may be of unequal size. It is not to be expected, however, that delay in realization will regularly reduce the size of sector by one-half for each cell generation. The sectors observed include only the cells of the aleurone layer which are included in the block of tissue affected; their reduction in size with advancing endosperm development depends upon the pattern of development of the tissues concerned. Actually, the average size of deficiency sector observed in ears treated at periods after pollination ranging from 73 to 81 hours was not changed enough to show any consistent relation among the 21 ears treated. Ears treated 96 hours after pollination yielded smaller but readily recognizable sectors. It is evident, therefore, that even if there were a delay in realization of gene mutation for several cell generations after irradiation, the resultant sectors would be detectable. Colored sectors as small as 10 cells in size are clearly recognizable on close examination of the seeds without magnification, and a single colored aleurone cell on a colorless seed is readily identified under magnification.

All of the colorless seeds were closely examined without magnification for sectors of colored aleurone. Occasional discolored areas of pericarp or endosperm were found, but under magnification these were readily distinguishable from colored aleurone sectors. No sectors of colored aleurone were found, among the 3832 seeds examined. On the basis of 1400 cells per seed, each containing 3 *a* genes capable of dominant mutation, this represents approximately 16 million chances for detectable dominant mutation (3832 × 1400 × 3). The treatment to which these cells were exposed produced in the adjoining colored seeds approximately 100,000 losses of *A*. Since the number of *A*'s tested for frequency of *A*-loss was only one-third the number of *a*'s tested for frequency of dominant mutation, the experiment shows failure to induce dominant mutation in a trial on a scale sufficient to produce about 300,000 *A*-losses.

A similar examination was made of the ears irradiated at 1600 r. At this dose the number of *A*-losses on seeds on the exposed side of the ear is too great to permit an accurate count. On the opposite side of the ear the frequency is very high also, but samples representative of the central portion of this region (amounting to approximately one-sixth of the ear) were suitable for examination. In 8 of the ears treated a sample of 6 seeds per ear was examined, the sample in each case being taken from the two rows least exposed to the radiation. The number of *A*-losses from these samples averaged 20.3 per seed. This is about three times the frequency of *A*-losses in seeds similarly located on the ears treated at 800 r, and inspection of seeds in other regions of the ears indicates that this is a fair estimate of the relative frequency of *A*-losses from the two doses used. The average size of sectors was 113 cells.

TABLE 1

FREQUENCY AND SIZE OF SECTORS DUE TO *A*-LOSS FOLLOWING X-RAY TREATMENT

Ear	No. 870	No. 876	No. 880	No. 881	No. 882	No. 884	Total
Interval, pollination to irradiation (hrs.)	75	75.5	80.5	81	81.5	75	
Number of sectors							
Seed No. 1	22	32	26	36	49	40	205
No. 2	30	31	38	33	34	29	195
No. 3	49	43	34	23	37	32	218
No. 4	3	9	13	14	6	15	60
No. 5	20	12	17	12	14	9	84
No. 6	11	16	13	18	24	22	104
Number of aleurone cells per sector (frequency distribution)							
1–10	12	12	16	9	23	18	90
11–50	38	43	48	42	47	41	259
51–200	71	64	46	63	82	69	395
201–500	14	23	27	22	11	19	116
500+	0	1	4	0	1	0	6

The 15 ears treated at 1600 r produced a total of 2487 colorless seeds. These were closely examined for sectors of colored aleurone, and in addition a sample of 50 seeds from the most heavily treated region of each was minutely examined under magnification. No sectors of colored aleurone were found, with the exception of a single doubtful spot possibly including 3 aleurone cells, which could not be positively identified. The frequency of *A*-losses, producible by the same dose in an equal population of *A*-genes, estimated as before, is about 600,000.

The susceptibility to dominant mutation of the *a* genes used in this experiment was tested by pollinating each of the parental stocks by a homozygous stock designated *a*-dl *Dt*. The gene *a*-dl is one of a number of "dotless" *a* alleles arising by mutation of standard *a* under the influence of *Dt*. It is indistinguishable from *a* in phenotypic effect, but differs in showing almost no dots due to dominant mutation in the presence of *Dt*. In endosperm triplex for *a*-dl and *Dt* an occasional small dot may be found, as in the case of the similar mutant a^s described by Rhoades, but the rate is negligible in comparison to that of standard *a*. In the endosperms of *a a a*-dl *dt dt Dt*, produced by the crosses under discussion, the dots observed may be ascribed almost wholly to mutation of one of the two *a*'s under the influence of the single dose of *Dt*. The mutation frequencies for x-ray action and *Dt* action are, of course, not directly comparable, since it is not possible to apply the two stimuli for comparable periods or to comparable populations of cells.

The number of dots per seed in endosperms of *a a a*-dl *dt dt Dt* produced by crosses on the female parental stock was 2.95, on the colorless seeds of the male parental stock 2.1. In each case the dots represent the number of mutations produced by two *a* genes in the presence of a single *Dt* gene. In the *a a a* seeds irradiated, therefore, the number of mutations to *A* which would be expected in the presence of one dose of *Dt* is about 4 per seed (2.95 + 2.1/2). Rhoades[4] has shown that the frequency of mutation of *a* to *A* is disproportionately increased by increased dosage of *Dt*, and that about 17 times as many mutations are produced in *Dt Dt Dt* seeds as in *Dt dt dt* seeds. The population of 6319 seeds of *a a a*, which failed to yield mutations to *A* under x-ray treatment, was therefore capable of yielding about 400,000 mutations to *A* under the influence of homozygous *Dt* (6319 × 4.0 × 17).

Summary.—X-ray treatment failed to induce mutation of *a* to *A* (or to any other colored aleurone allele), in populations which, with the x-ray doses applied, were capable of yielding about 900,000 losses of *A* by deficiency or by mutation to a colorless allele.

The *a* genes used in the experiment, when combined with *Dt*, yielded numerous mutations to *A* or other colored aleurone alleles. The population irradiated was large enough to have yielded, under the influence of homozygous *Dt*, about 400,000 of such mutations.

---

[1] Coöperative investigations of the Division of Cereal Crops and Diseases, Bureau of Plant Industry, Soils, and Agricultural Engineering, Agricultural Research Administration, U. S. Department of Agriculture, and Missouri Agricultural Experiment Station, University of Missouri. Missouri Agricultural Experiment Station Journal Series 940.

[2] Stadler, L. J., *Proc. Sixth Int. Cong. Genetics*, **1**, 274–294 (1932).

[3] Stadler, L. J., *Cold Spring Harbor Symposia*, **9**, 168–177 (1941).

[4] Demerec, M., these PROCEEDINGS, **20**, 28–31 (1934).

[5] Rhoades, M. M., *Genetics*, **23**, 377–395 (1938).

[6] Rhoades, M. M., *Cold Spring Harbor Symposia*, **9**, 138–144 (1941).

GRAPHICAL METHOD OF FACTORING THE CORRELATION MATRIX

BY L. L. THURSTONE

THE PSYCHOMETRIC LABORATORY, THE UNIVERSITY OF CHICAGO

Communicated April 14, 1944

Multiple factor analysis starts with a square matrix R of order $n \times n$ with cell entries r_{jk} which are experimentally determined. The matrix R shows the correlation of each variable with every other variable in the test battery. The correlation matrix is symmetric since $r_{jk} = r_{kj}$. Some students write unity in the diagonal cells and this is legitimate for some problems. More often the diagonals contain the communalities h_j^2 which are so chosen as to be consistent with the rank of the matrix as determined by the side entries. The factoring methods to be described can be applied for any set of diagonal entries.[1]

The first object of multiple factor analysis is to find a factor matrix F of order $n \times r$ where r is the rank of the correlation matrix R_{jk}. The factor matrix F must be written so as to satisfy the fundamental relation $FF' = R$. The correlation matrix R can be factored into a factor matrix F in many different ways and F is not here unique. The current geometrical interpretation of R is that its entries r_{jk} show the scalar products of pairs of test vectors J and K which are not necessarily of unit length unless this restriction is imposed by unit diagonals r_{jj}. Since the correlation coefficients are scalar products, it follows that no reference frame is implied by R. Each row j of the factor matrix F shows the entries a_{jm} which are the projections of the test vector J on a set of r orthogonal reference axes. The lack of uniqueness of F is geometrically interpreted in the free choice of an orthogonal reference frame for the test configuration of n test vectors J. The reference frame is explicitly defined in F relative to the test vectors but it is not defined in R.

One factorial solution for F is to locate the first reference axis so as to maximize the sum of squares of test vector projections. The second reference axis can be located by the same criterion applied to first factor correlation residuals, and so on until a set of r orthogonal axes have been determined which correspond to rank r of the correlation matrix. In the writer's original formulation of this problem, it was called the *principal axes solution.*[2] In the first paper on this problem a computational method was described for finding the principal axes solution F. It involved first factoring the correlation matrix into a factor matrix F with any arbitrary orthogonal reference frame and a subsequent orthogonal transformation representing a rigid rotation to the principal axes. That solution involved the characteristic equation of the matrix $F'F$ and determination of its latent roots. The next year Hotelling described an iterative method of factoring the correlation matrix directly into the principal axes solution which he renamed *principal components*.[8] Since the iterative methods, as well as all other available methods, of obtaining the principal axes solution are very laborious, it is of interest to students of multiple factor analysis to find shorter methods.

Consider the first column of F for the principal axes solution. Its elements are a_{j1}. It is desired to find the values of a_{j1} so that the first factor shall account for a maximum of the variance in the correlation matrix R. Then

$$r_{jk} - a_{j1}a_{k1} = \rho_{jk}, \tag{1}$$

where a_{j1} and a_{k1} are the first factor saturations of tests j and k, and ρ_{jk} is the residual to be minimized. Squaring and summing for all correlation coefficients, we have

$$\sum_j\sum_k r_{jk} - 2\sum_j\sum_k\Sigma a_{j1}a_{k1}r_{jk} + \sum_j\sum_k a_{j1}^2a_{k1}^2 = \sum_j\sum_k \rho_{jk}^2 \equiv z, \tag{2}$$

where the sum of the squares of the correlation residuals ρ_{jk} may be denoted z. The partial derivative with respect to a_{j1} is then

$$\frac{\partial z}{\partial a_{j1}} = -4\sum_k r_{jk}a_{k1} + 4a_{j1}\sum_k a_{k1}^2 \tag{3}$$

from which we get

$$a_{j1}\sum_k a_{k1}^2 = \sum_k r_{jk}a_{k1}. \tag{4}$$

Let the sum $\Sigma a^2 \equiv t$. The summation Σra can be written as a matrix product Ra. Then

$$Ra = at \tag{5}$$

where R is the given square matrix, a is a column vector, and t is a scalar, namely, the largest latent root. The iterative process consists in starting with a trial vector u so that

$$Ru = v\Sigma u^2 \tag{6}$$

where $(v\Sigma u^2)$, or numbers proportional to these, constitute the trial vector u for the next iteration. The process continues until $v\Sigma u^2$ becomes proportional to the trial values u at which time the values of v $(= u)$ are the desired values of a_{j1} and the first column of F is then determined. The number of iterations can be considerably reduced by starting with R^2, or R^4, or even R^8. While that looks effective in a text book example, it is not so inviting a procedure when we are confronted with a correlation matrix of order 70 × 70.

Consider the matrix multiplication Ru in equation (6) and, in particular, the scalar product of the row j of R_{jk} and the column vector u. This product is

$$R_j u = v_j \Sigma u^2 \qquad (7)$$

and this can be written briefly as

$$v_j = \frac{R_j u}{\Sigma u^2} \qquad (8)$$

but this is also the well-known equation for the slope of the regression line r on u through the origin for a plot of r against u. If we plot columns of r_{jk} against the trial values u_k, we get a plot with n points. The best fitting straight line through the origin (regression r on u) can be drawn by inspection and the slope of the line is easily read graphically with sufficient accuracy for the first few iterations. That slope is the value v_j. To plot, say, 30 points and to draw by inspection the best fitting straight line through the origin takes less time than to add cumulatively 30 products on a calculating machine.

By equations (5) and (6) it is evident that at the solution the values $v_j = u_j = a_j$. When a set of values of v has been determined graphically, we plot v against u and find the slope m. In the first few graphical iterations, this slope will not be unity because the trial values u are probably too large or too small. If the true values are better represented by ku, in which k is a stretching factor, then a plot of r against ku will give a slope of v/k instead of v. When these graphs have been drawn, the slope of v/k against ku should be unity. But the obtained slope of v against u is m for the values actually used. Hence the observed slope $m = k^2$. Having found the slope m of the plot of v against u, we find $k = \sqrt{m}$. Then if we should take a new set of trial numbers $x_j = ku_j$, the iteration would give, instead of equation (8),

$$\frac{Rx}{\Sigma x^2} = y, \qquad (9)$$

where $y = v/k$. The slope of y against x should now be unity and the values of y should be used as the trial vector for the next graphical iteration. The determination of m and k can be done with a slide rule and the new trial values $y = v/k$ can also be determined by the slide rule. The slope m can be found either by inspection or by simple summing with the method of averages in which $m = \Sigma v/\Sigma u$ for like signed pairs of v and u.

The graphical method here described for the principal axes solution has also been adapted for a centroid method of factoring the correlation matrix by which the computations for a factor matrix can be done by a computer without excessive labor even for a large matrix and without the use of tabulating machine equipment. That procedure is considerably faster and it will be described in a subsequent paper.

A numerical example is here given. In table 1 we have a correlation matrix R of order 10×10 and rank 3. The corresponding factor matrix F for the principal axes is shown in the same table. It is the solution to be found from R. The latent roots of R are 2.85, 1.47 and 0.63. These are the sums of squares of columns of F. In table 2 the initial trial vector u for the major principal factor is taken roughly proportional to the sums of columns of R. Column 1 of R is plotted against column u and the slope is estimated to be approximately $+0.05$ and it is recorded in column v_1. Ten such plots give the values in column v_1. The slope is read directly from each graph by noting the ordinate of a straight line fit at $u_1 = 1$. Next obtain the two column sums Σu_1 and Σv_1. The ratio $\Sigma v/\Sigma u = m$ and $k = \sqrt{m}$. This summation can be absolute sums for like signed pairs of u and v. Then compute $y_1 = v_1/k$ with a slide rule. These are also the trial values u_2 for the next trial. Proceed likewise for three trials which give the desired values of a_{j1} to two decimals.

The first factor residuals are then computed. These are $\rho_{jk} = r_{jk} - a_{j1}a_{k1}$. Choose as a starting vector that column of the first factor residuals which has the largest absolute sum, ignoring diagonals. This is column 1. Its values are recorded in column u_1 for factor two. The procedure is now the same as before. The third iteration was taken on a calculating machine and it gave the second factor loadings a_{j2}, the second column of F, to two decimals. The last iterations for each factor can be done on a calculating machine to obtain greater accuracy while the first few iterations can be done graphically to save time and labor. The computation for the third column of F was done in the same manner.

Occasionally, when two or three latent roots are nearly the same, the iterative process will be found to oscillate or the convergence will be slow. This is an indication that the test configuration has nearly equal thicknesses in these dimensions, so that the configuration is nearly circular or spherical in these dimensions. If the purpose is to extract the maximum variance from the correlation matrix, it does not matter where we place a set of orthogonal axes in those dimensions which are represented by nearly equal latent roots. In such a situation it is desirable to reduce the obtained vector v to a unit vector in the system without waiting for complete convergence. That can be done with a stretching factor p on v so that $pv = a$, where a are the desired factor loadings. The value of p can be found from the relation $1/p = \sqrt{\Sigma uv}$. By this device the computer need not be unduly delayed by slow convergence in obtaining a factorially useful solution.

TABLE 1

	CORRELATION MATRIX R									
	1	2	3	4	5	6	7	8	9	10
1	0.37	0.33	−0.29	0.11	0.10	0.16	−0.33	−0.09	0.20	−0.03
2	0.33	0.46	−0.31	0.18	0.22	0.05	−0.15	−0.14	0.20	0.18
3	−0.29	−0.31	0.30	−0.05	−0.04	0.04	0.34	0.27	0.03	0.15
4	0.11	0.18	−0.05	0.14	0.18	0.15	0.11	0.14	0.28	0.30
5	0.10	0.22	−0.04	0.18	0.24	0.16	0.20	0.18	0.34	0.42
6	0.16	0.05	0.04	0.15	0.16	0.45	0.08	0.43	0.54	0.33
7	−0.33	−0.15	0.34	0.11	0.20	0.08	0.70	0.43	0.27	0.63
8	−0.09	−0.14	0.27	0.14	0.18	0.43	0.43	0.62	0.56	0.54
9	0.20	0.20	0.03	0.28	0.34	0.54	0.27	0.56	0.77	0.66
10	−0.03	0.18	0.15	0.30	0.42	0.33	0.63	0.54	0.66	0.90
Σ	0.53	1.02	0.44	1.54	2.00	2.39	2.28	2.94	3.85	4.08

	FACTOR MATRIX F		
	I	II	III
1	0.00	0.60	0.10
2	0.10	0.60	−0.30
3	0.20	−0.50	0.10
4	0.30	0.20	−0.10
5	0.40	0.20	−0.20
6	0.50	0.20	0.40
7	0.60	−0.50	−0.30
8	0.70	−0.20	0.30
9	0.80	0.30	0.20
10	0.90	0.00	−0.30
Σ	4.50	0.90	−0.10

TABLE 2

	FACTOR I						FACTOR II					
	TRIAL 1		TRIAL 2		TRIAL 3		TRIAL 1		TRIAL 2		TRIAL 3	
	u_1	v_1	$y_1 = u_2$	v_2	$y_2 = u_3$	v_3	u_1	v_1	$y_1 = u_2$	v_2	$y_2 = u_3$	v_3
1	0.10	0.05	0.05	0.10	0.10	0.00	0.37	1.00	0.59	0.60	0.61	0.60
2	0.25	0.15	0.16	0.10	0.10	0.10	0.33	1.00	0.59	0.62	0.63	0.60
3	0.10	0.15	0.16	0.20	0.20	0.20	−0.29	−0.80	−0.47	−0.50	−0.51	−0.50
4	0.40	0.30	0.33	0.30	0.30	0.30	0.11	0.35	0.21	0.15	0.15	0.20
5	0.50	0.40	0.44	0.40	0.40	0.40	0.10	0.35	0.21	0.20	0.20	0.20
6	0.60	0.50	0.55	0.50	0.51	0.50	0.16	0.50	0.30	0.20	0.20	0.20
7	0.60	0.50	0.55	0.60	0.61	0.60	−0.33	−0.90	−0.53	−0.50	−0.51	−0.50
8	0.70	0.60	0.66	0.70	0.71	0.70	−0.09	−0.30	−0.18	−0.20	−0.20	−0.20
9	0.90	0.80	0.87	0.80	0.81	0.80	0.20	0.55	0.33	0.30	0.31	0.30
10	1.00	0.85	0.93	0.90	0.91	0.90	−0.03	0.00	0.00	0.00	0.00	0.00
Σ	5.15	4.30	4.70	4.60			2.01	5.75	3.41	3.27		
m		0.8350		0.9787				2.860		0.959		
k		0.9138		0.9893				1.691		0.979		

The graphical method here described for estimating inner products can be applied in a variety of computational problems.

---

[1] The author wishes to express his indebtedness to the Carnegie Corporation of New York for a research grant in support of our development of multiple factor analysis and its application to the study of primary mental abilities.

[2] L. L. Thurstone, "Theory of Multiple Factors," 1932, pp. 17–27, Edward Brothers, Ann Arbor, Mich.

[3] Harold Hotelling, "Analysis of a complex of statistical variables into principle components," *J. Educ. Psych.*, **24**, 417–441, 498–520, (1933).

THE DEVELOPMENT OF NORMAL AND HOMOZYGOUS BRACHY (T/T) MOUSE EMBRYOS IN THE EXTRAEMBRYONIC COELOM OF THE CHICK†

BY S. GLUECKSOHN-SCHOENHEIMER*

DEPARTMENT OF ZOÖLOGY, COLUMBIA UNIVERSITY

Communicated April 15, 1944

The study of the causal morphology of mammalian embryos in spite of the great interest attached to it has not progressed very far because of the technical difficulties with which any experimental approach to the problem has met. Waddington and Waterman[1] grew rabbit embryos *in vitro* for a limited time and Nicholas and Rudnick[2] devised a method for raising rat embryos in a culture medium. In spite of these attempts, it was not possible to operate on mammalian embryos in early stages and let them continue their development inside the uterus, nor could such embryos be raised and develop normally in a suitable extra-uterine medium for any length of time. Methods which would accomplish this would be of importance not only for the study of normal causal morphology of mammalian embryos, but also for an experimental study of the embryogeny of certain hereditary abnormalities. Searching for such a method, a new procedure for raising entire mouse embryos outside the uterus was developed and described in detail (Gluecksohn-Schoenheimer).[3] It consisted in removing the embryos from the uterus of the mother and transplanting them into the extra-embryonic coelom of the chick embryo where they remained and developed for one or several days. This method has been applied in the experiments reported here to the study of embryos homozygous for the Brachyury mutation *TT*.

As described by Chesley[4] the homozygous mutants (*T*/*T*) show extreme morphological abnormalities from the age of about 8 days on and die at about 10 days. The posterior body region, including posterior limb buds, is missing completely and extensive abnormalities are found in notochord, neural tube and somites.

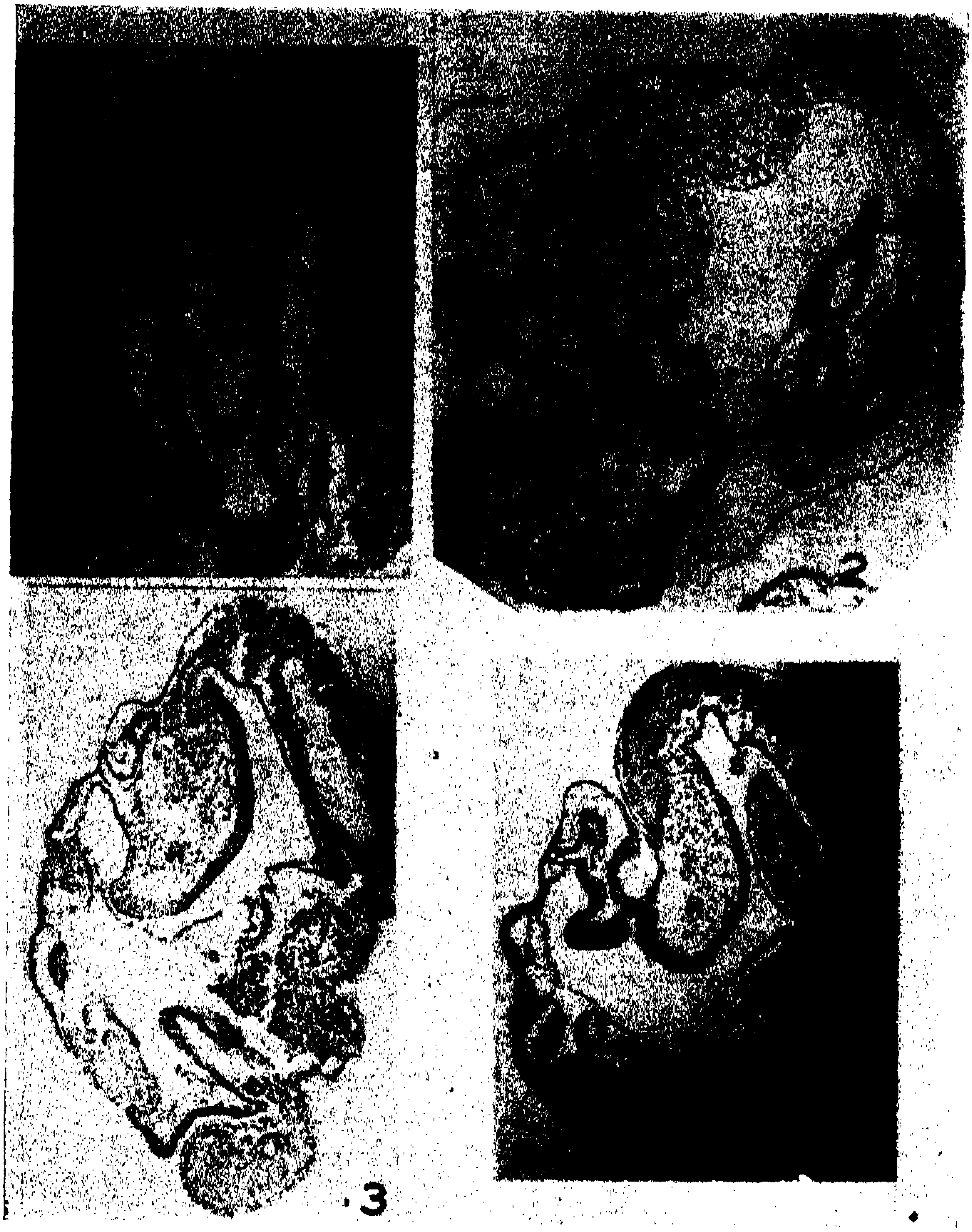
An additional abnormality, not mentioned by Chesley, was observed in the allantois. While the normal allantois, after having budded off from the embryo, appears turgid and proceeds to grow and to lengthen until it comes in contact with the ectoplacental cone, the abnormal allantois of the T/T homozygotes stays very short and has a shrunken tuberous appearance. As a result of the abnormality of the allantois, umbilical vessels do not develop; thus the connection of the embryonic and the maternal circulation fails to be established and the embryo dies at the time when maternal circulation becomes all-important for the nutrition of the embryo (age of 10 days).

Ephrussi,[5] in an attempt to determine the potentialities of the different tissues of the inviable T/T embryos, explanted these tissues and grew them in tissue culture. Under the experimental conditions, the tissues survived, grew and differentiated beyond the time and stage at which they would have died *in utero*. From his results, Ephrussi concludes that the death of the embryos *in utero* is probably caused by a "disturbance of the normal correlations" and is due "exclusively to internal factors." From our observations, it seems that the death of the embryo at the particular time of 10 days after fertilization is due to the failure of the circulatory connection with the mother rather than to any unknown internal factors.

It seemed of considerable interest to study the extra-uterine development of whole T/T embryos from two points of view: first, it was planned to examine their development in the chick host and compare it with that of embryos of the same genotype (T/T) in their normal environment; secondly, it was hoped that by this approach we might be able to learn something about the rôle of the allantois in the development of the homozygotes.

For this purpose, embryos derived from matings of heterozygous Brachy females ($T/+$) to heterozygous Brachy males ($T/+$) were used. A total of 149 transplants was made, 84 of which were successful. A transplant was considered successful if development of the mouse embryo proceeded after transplantation and if the embryo was alive when removed from the chick host. The mouse embryos were left in the chick embryo for 18 to 47 hours. After this time, they were taken out from the chick host, examined, recorded, fixed in Bouin, imbedded in paraffine, sectioned at 6 μ and stained in Delafield's Haematoxylin and eosin. The results to be reported here were arrived at after examination of the sections.

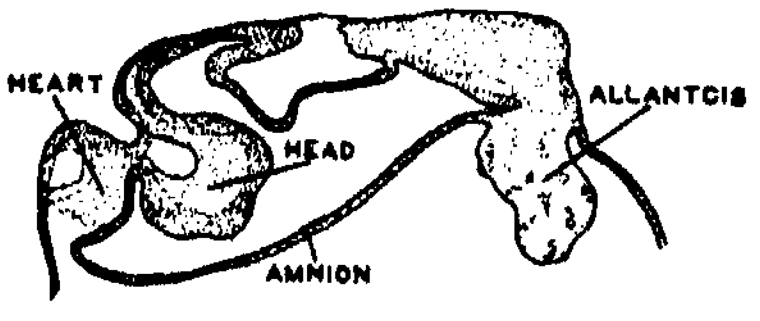
As described before (Gluecksohn-Schoenheimer),[3] normal mouse embryos removed from the uterus of the mother 7 days after copulation (egg cylinder stage, Fig. 1) and transplanted into the extra-embryonic coelom of the chick develop and differentiate head structures, neural folds, somites, a beating heart and a tail bud (Fig. 2). The embryos from matings of Brachy ($T/+$) by Brachy ($T/+$), transplanted into the extra-



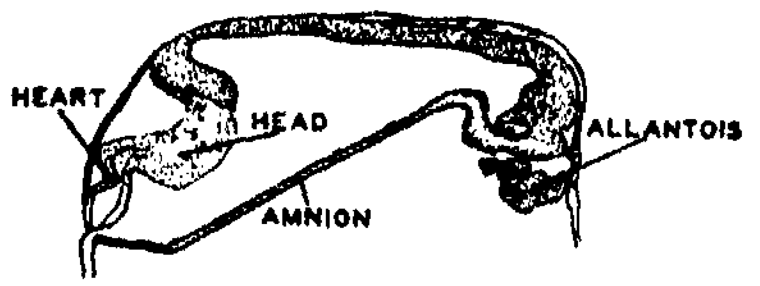
FIGURES 1-4

Figure 1. Sagittal section through an egg cylinder stage of a normal mouse embryo, 7 days 19 hours after copulation, = stage 11. *A* = Allantois. 80×. Figure 2. Sagittal section through a normal mouse embryo (No. 207) explanted into the extra-embryonic coelom of the chick in the egg cylinder stage (stage 11) and removed from there 49 hours later. *H* = head, *h* = heart. 80×. Figure 3. Sagittal section through a mouse embryo (No. 314) from a mating of *T*/+ × *T*/+, explanted into the extra-embryonic coelom of the chick in the egg cylinder stage (stage 11) and removed from there 24 hours later. Embryo belongs into Class I = normal. *H* = head, *S* = somites. 80×. Figure 4. Sagittal section through the same embryo (No. 314). Note turgid attached allantois. *H* = head, *S* = somites, *A* = allantois. 80×.

embryonic coelom of the chick, fall into three classes: the first class comprises those embryos which develop normally in the host environment; they show good development of head, heart, somites and neural tube (Fig. 3); the allantois is big and turgid and attached to the ectoplacental cone (Fig. 4). The rest of the embryos show abnormal development and can, in turn, be separated into two classes, using the condition of the allantois as criterion. In both of these classes, II and III, the allantois is abnormal, but in each class typically different from the other: Class II contains embryos with a short allantois which does not attach to the ectoplacental cone; but while it is not big and turgid, it is not shrunken either. It looks as though it had been arrested in its development towards normal. The general condition and development of the embryos in this class are not as good as of those in the normal group: head and trunk structures are underdeveloped and somites are sometimes lacking entirely (Fig. 5). Whether there exists any causal relationship between the underdevelopment of the allantois and that of the rest of the embryo, or whether the abnormal allantois presents just one symptom of a general tendency to underdevelopment of the entire embryo cannot be decided now. The embryos of the



5



6

7

FIGURES 5–7

Figure 5. Sagittal section through mouse embryo (No. 413) from mating of *T*/+ × *T*/+, grown in extra-embryonic coelom of the chick; Class II. No somites, underdevelopment of head and trunk. Short allantois. 80×. Drawn from photograph. Figure 6. Sagittal section through mouse embryo (No. 312) from mating of *T*/+ × *T*/+, grown in extra-embryonic coelom of chick; Class III. No somites, underdevelopment of head and neural folds; note the short tuberous allantois. 80×. Drawn from photograph. Figure 7. Photograph of a mouse embryo (*T*/*T*, Class III) which was explanted into the extra-embryonic coelom of the chick in the early head fold stage (8 days) and remained there for 47 hours. Typical features of *T*/*T* embryo: development of head, heart and anterior limb-buds, absence of posterior body structures. Compare Chesley[4], Fig. 6. 20×.

second abnormal group, Class III, are characterized by an allantois with a very typical abnormal appearance: it is small, not turgid but collapsed, and characteristically tuberous; it stays short and never attaches to the ectoplacental cone. Its most characteristic feature is its tuberosity (Fig. 6). There is definite underdevelopment in the rest of the embryo; the head is small, the trunk short, the neural folds are thin, there are no somites, and hardly ever a beating heart. Four of the embryos in Class III developed into typical *T/T* homozygotes (Fig. 7); all of these had been explanted in early head fold stages, i.e., at just 8 days after copulation. The rest of the embryos in this class which had been explanted in earlier stages—i.e., egg cylinder stage—could not be diagnosed by any of the typical features of 9-day *T/T* homozygotes because they did not develop far enough. But they were diagnosed by the typically abnormal condition of the allantois; and in addition they showed underdevelopment of head and trunk and absence of somites.

TABLE 1

	EMBRYOS OBTAINED				
	CLASS I NORMAL	CLASS II ABNORMAL	CLASS III ABNORMAL *T/T*	?	TOTAL
Explantation stage of embryos,					
Stage: 7–10[a]	3	4	2		9
Stage: 11–13[b]	8	22	11	5	46
Stage: early head fold, 8 days	17	1	6	2	26
Total	28	27	19	7	81

[a] Early egg cylinder stages, end of 6 days.

[b] Egg cylinder stages, 7 days, beginning growth of allantois, slight thickening of head folds.

When we consider the numerical distribution of the explanted embryos among the three different classes (table 1), we find that 19 out of a total of 81 fall into Class III, i.e., the group comprising the embryos supposed to represent the abnormal *TT* class. Of these 19, only 4 developed all of the typical *TT* features; but all of the others had the typically abnormal tuberous allantois which is found in all *TT* embryos developing in their normal environment (see above) and which never appears in explanted embryos from matings of normal by normal. We expect one quarter of the embryos from *T*/+ by *T*/+ matings to be *T/T* and the actual number of 19 out of 81 fits expectation very well.

The second class of abnormals—27 embryos—might possibly represent the *T*/+ embryos. If one considers the different explantation stages separately in this class, it appears that of the embryos explanted at earlier stages (7–13) about 50% fall into Class II while of those explanted at later stages (early head fold and 8-days stage) about 70% fall into Class I and only one embryo is classified as Class II. It is conceivable that the developmental potencies of the very early $T/+$ embryos under the unfavorable explantation conditions are reduced in comparison with those of $+/+$ embryos. If this is the case, it might well be that the allantois, which is completely abnormal in the homozygotes, reacts to the abnormal explantation conditions more strongly in the young heterozygous embryo than it does in the homozygous normal, showing lower developmental potencies under abnormal conditions than does the normal. In this connection it seems of interest that a deficiency of the number of heterozygous short-tailed mice ($T/+$) appeared in matings of Brachy short-tailed mice ($T/+$) *inter se* (Kobozieff and Pomriaskinsky-Kobozieff)[6] and in our own laboratory (Dunn).[7] Although some of this deficiency could be accounted for by the failure of the Brachy mutation to express itself in some of the heterozygotes, a differential prenatal mortality of $T/+$ embryos might be responsible for the lower number of heterozygotes in addition to the normal appearance of some $T/+$ mice. The total number of embryos in Class II does not correspond to the expected 50% of $T/+$ embryos while the number of 28 in the normal group is higher than the expected 25%. One would thus have to assume that a certain fraction of $T/+$ embryos, namely, the older ones, were able to develop normally, or one might ascribe the distribution of embryos among the two classes entirely to chance. In the latter case, one would assume identical developmental potencies for $+/+$ and $T/+$ embryos, and they would be equally distributed between Classes I and II. Their total of 55 corresponds well to the expected 75% of $+/+$ and $T/+$ embryos combined. While, thus, there remain some doubts about ascribing the embryos in Classes I and II to two genetically different groups, $+/+$ and $T/+$, the remaining 25% fall into a group distinctly and characteristically different from the rest of the embryos, independent of the stage of explantation; this group is assumed to represent the T/T class.

It thus appears that the abnormal development of T/T embryos is determined before it is apparent morphologically and is independent of the stage of explantation (table 1). The earliest stages which it is possible to explant, i.e., early egg cylinder stages, show the tendencies to abnormal development of the entire embryo. Earlier stages cannot be tested, since they do not develop at all under the explantation conditions described, regardless of whether they come from crosses of normals or Brachys.

An analysis of the rôle of the allantois in the development of the normal and abnormal embryos was not possible on the basis of our experiments. The fact that an abnormal or underdeveloped allantois occurs always in combination with abnormal or subnormal development of the entire embryo seems highly suggestive. The allantois of the mouse might well serve as an organ for intake of nutrient substances and excretion of waste products. Any disturbance in the function of this organ might thus lead to secondary abnormalities throughout the entire embryo. However, experiments involving the extirpation of the allantois and studying its effects on the embryo would be required to settle the question of the rôle of the allantois in the development of the normal mouse embryo. Such experiments might also provide an answer to the question of the possible causal rôle of the allantois in the abnormal development of *TT* embryos.

Summary.—Embryos from crosses of Brachy short-tailed mice ($T/+$) *inter se* were removed from the mother in early stages of development and transplanted to the extra-embryonic coelom of the chick. They remained in the host for 18 to 47 hours and were then removed for study. The embryos obtained could be separated into three different classes: (1) normal embryos; (2) abnormal embryos with different degrees of abnormalities, with reduced developmental potencies, possibly representing the $T/+$ embryos; (3) abnormal embryos with typical abnormalities, representing T/T embryos. T/T embryos will thus develop true to their genotype, even if explanted before the time at which morphological differences appear. The possible causal rôle of the allantois in the development of these embryos is discussed.

† This research was aided by grants from the Josiah Macy, Jr., Foundation and the Fund for Research of Columbia University.

* The author wishes to express her gratitude to Dr. L. C. Dunn for his suggestions and his constant interest during the course of this work.

[1] Waddington, C. H., and Waterman, A. J., *Jour. Anat.*, **67**, 355–370 (1933).

[2] Nicholas, T. S., and Rudnick, D., *Jour. Exp. Zool.*, **78**, 205–232 (1938).

[3] Gluecksohn-Schoenheimer, S., *Science*, **93**, 502–503 (1941).

[4] Chesley, P., *Jour. Exp. Zool.*, **70**, 429–459 (1935).

[5] Ephrussi, B., *Ibid.*, **70**, 197–204 (1935).

[6] Kobozieff, N., and Pomriaskinsky-Kobozieff, N. A., *Compt. rend. des séances de la Société de biologie*, **127**, 120 (1938).

[7] Dunn, L. C., *The Harvey Lectures*, **35**, 135–165 (1940).

REMARK ON A DISTORTION TENSOR FOR ELASTIC DISPLACEMENTS

By T. Y. Thomas

University of California at Los Angeles and Institute for Advanced Study

Communicated April 18, 1944

One of the most successful of the various conditions for yield resulting from a state of stress produced by mechanical means is that due to von Mises.[1] This condition states that yield will occur when the energy per unit volume produced by a change of shape alone reaches a certain value characteristic of the material under consideration. It would thus appear that distortion, i.e., change of shape as distinct from change of volume, is one of the primary factors in the occurrence of yield. This suggests that in addition to the ordinary tensors of stress and strain (related by Hooke's law) we consider a distortion tensor and seek to express the condition of yield directly in terms of this tensor.

A differentiable point transformation $x \rightarrow \bar{x}$ referred to rectangular coördinates and having its jacobian everywhere different from zero (and hence locally non-singular) will be conformal (i.e., angles will be left unaltered) if, and only if,

$$\Sigma \frac{\partial \bar{x}^\alpha}{\partial x^\beta} \frac{\partial \bar{x}^\alpha}{\partial x^\gamma} = \rho \delta_{\beta\gamma}, \qquad (\rho > 0), \tag{1}$$

where ρ is some positive function of position. When the transformation is an elastic displacement it is customarily represented by $\bar{x}^\alpha = x^\alpha + u^\alpha(x)$ in which the u^α (also denoted by u_α since rectangular coördinates are employed) are the components of a vector. In accordance with the usual assumptions, the u's and their derivatives are small quantities, the product of which can be neglected. Hence

$$\frac{\partial \bar{x}^\alpha}{\partial x^\beta} = \delta^\alpha_\beta + u^\alpha_{,\beta}, \tag{2}$$

and

$$\delta_{\beta\gamma} + u_{\beta,\gamma} + u_{\gamma,\beta} = \rho \delta_{\beta\gamma}, \tag{3}$$

when (2) is substituted into (1). Taking the determinant of both members of (1) and making use of the form (2) of the derivatives $\partial \bar{x}^\alpha / \partial x^\beta$ we have

$$|\delta^\alpha_\beta + u^\alpha_{,\beta}|^2 = |\rho \delta_{\alpha\beta}| = \rho^3,$$

or

$$\rho = |\delta^\alpha_\beta + u^\alpha_{,\beta}|^{2/3} = 1 + {}^2/_3 u^\alpha_{,\alpha} = 1 + {}^2/_3 \Delta, \qquad (\Delta = u^\alpha_{,\alpha}).$$

Substitution of this value of ρ into (3) gives

$$e_{\alpha\beta} - \frac{\Delta}{3} \delta_{\alpha\beta} = 0, \qquad (e_{\alpha\beta} = {}^1/_2 (u_{\alpha,\beta} + u_{\beta,\alpha})), \tag{4}$$

where e is the strain tensor of elasticity theory. The quantity Δ is the dilatation or increase per unit volume due to the displacement. Conversely if (4) is satisfied the elastic displacement is conformal.

We now define a tensor D as an invariant of an arbitrary elastic displacement by putting

$$D_{\alpha\beta} = e_{\alpha\beta} - \frac{\Delta}{3} \delta_{\alpha\beta}.$$

We call D the *distortion tensor*. From the above, the displacement will be conformal (i.e., without distortion) if, and only if, $D = 0$. We can therefore think of the tensor D, in general, as giving a measure of the distortion produced by the elastic displacement.[2]

It is well known that any elastic displacement causes a displacement of the local neighborhood of an arbitrary point P which is the result of the following three displacements, namely (α) a translation $P \rightarrow Q$, (β) a rigid rotation about Q, and (γ) a non-rigid displacement or deformation having Q as a fixed point. This latter deformation can be resolved, on the basis of the distortion tensor D, into a component with distortion and a component without distortion. Thus if δx denotes an arbitrary small vector issuing from the above point Q, the part of the deformation due to the distortion will be represented by $D_{\alpha\beta}\delta x^{\beta}$ and the part without distortion by $(\Delta/3)\delta_{\alpha\beta}\delta x^{\beta}$ these quantities being the components of the changes produced in δx^{α} in this decomposition of the deformation.

Hooke's law for a homogeneous and isotropic medium can now be written

$$E_{\alpha\beta} = 2\mu D_{\alpha\beta}, \qquad \Theta = (3\lambda + 2\mu)\,\Delta, \tag{5}$$

where

$$E_{\alpha\beta} = \sigma_{\alpha\beta} - \frac{\Theta}{3}\delta_{\alpha\beta}, \qquad \Theta = \Sigma\sigma_{\alpha\alpha}.$$

Thus E is the distortion analogue of the stress tensor σ while λ and μ are the usual elastic constants. Now denote by τ_i and γ_i ($i = 1, 2, 3$) the principal values (at any point) of the tensors σ and E, respectively. From the definition of these quantities we immediately have the relations $\gamma_i = \tau_i - \Theta/3$ and when written in full these become

$$\left.\begin{aligned} \gamma_1 &= {}^1/_3[(\tau_1 - \tau_2) + (\tau_1 - \tau_3)], \\ \gamma_2 &= {}^1/_3[(\tau_2 - \tau_3) - (\tau_1 - \tau_2)], \\ \gamma_3 &= {}^1/_3[-(\tau_1 - \tau_3) - (\tau_2 - \tau_3)]. \end{aligned}\right\} \tag{6}$$

Now consider the square of the tensor E. Denoting this by E^2 we have

$$\begin{aligned} E^2 &= \Sigma E_{\alpha\beta}E_{\alpha\beta} = \Sigma(\gamma_\alpha\delta_{\alpha\beta})(\gamma_\alpha\delta_{\alpha\beta}) = \Sigma\gamma_i\gamma_i \\ &= {}^1/_3[(\tau_1 - \tau_2)^2 + (\tau_1 - \tau_3)^2 + (\tau_2 - \tau_3)^2], \end{aligned}$$

when use is made of the above relations (6). But apart from a constant factor the above bracket expression is the energy per unit volume due to distortion.[3] Since the squares of the tensors D and E differ only by a constant factor on account of (5), the von Mises yield condition can therefore be stated as follows: *yield will occur when the square of the distortion tensor D (or its stress analogue E) attains a characteristic value.*

By a distortion invariant of an elastic displacement we understand any invariant which depends solely on the distortion. In particular any scalar function of the components $\sigma_{\alpha\beta}$ (or $e_{\alpha\beta}$) will be a distortion invariant if, and only if, it is expressible as a function of the components $E_{\alpha\beta}$ (or $D_{\alpha\beta}$). The sufficiency of this condition is immediately obvious. To show the necessity for a scalar $S(\sigma_{\alpha\beta})$ we have

$$S(\sigma_{\alpha\beta}) = S(\tau_\alpha \delta_{\alpha\beta}) = S[(\tau_\alpha + \tau)\delta_{\alpha\beta}]$$

where τ is arbitrary. But choosing $\tau = -\Theta/3$ the latter function is seen to become $S(\gamma_\alpha \delta_{\alpha\beta})$or $S(E_{\alpha\beta})$ from which the conclusion follows. This result can obviously be extended to differential and non-scalar invariants. The above quantity E^2 or D^2 is an example of a distortion scalar. So likewise is the quantity $\gamma_1 - \gamma_3 = \tau_1 - \tau_3$ which is proportional to the maximum shearing stress (when $\tau_1 \geqq \tau_2 \geqq \tau_3$). This stress also appears to be of significance in determining the condition of yield.[4] Other examples are the magnitudes of the principal values γ_i of the tensor E (which differ from the principal values of the distortion tensor by a constant factor). Useful modifications of the known yield conditions may possibly be given by these latter or other distortion scalars which can be constructed. Such invariants will automatically take care of the failure to reach a state of yield as the result of hydrostatic pressure or uniform tension.

---

[1] *Nachr. d. Gesellsch. d. Wissensch. zu Göttingen, Math.-Phys. Klasse*, pp. 582–592 (1913). See also, Nadai, A., *Plasticity*, McGraw-Hill, Chap. 13 (1931), and Timoshenko, S., *Theory of Elasticity*, McGraw-Hill (1934), p. 137.

[2] The usual formula for the change $\delta\theta$ of an angle θ due to an elastic displacement gives this change in terms of the components $e_{\alpha\beta}$ of the strain tensor. Cf. Love, *The Mathematical Theory of Elasticity*, Cambridge, Vol. 4 (1927), p. 62. Also, Timoshenko, S., loc. cit., p. 192. By making use of the above definition of the components $D_{\alpha\beta}$ it can be shown that $\delta\theta$ is given by

$$\sin\theta\ \delta\theta = (D_1 + D_2)\cos\theta - 2D_{\alpha\beta}\nu_1^\alpha \nu_2^\beta$$

where,

$$D_1 = D_{\alpha\beta}\nu_1^\alpha \nu_1^\beta, \qquad D_2 = D_{\alpha\beta}\nu_2^\alpha \nu_2^\beta,$$

and ν_1 and ν_2 are unit vectors determining the angle θ. Hence the change in an angle produced by an elastic displacement is expressible directly in terms of the distortion tensor D.

[3] See Nadai, S., loc. cit., p. 72, and Timoshenko, S., loc. cit., p. 137.

[4] See, for example, Nadai, A., loc. cit., p. 60.

THE MINIMUM RANK OF A CORRELATION MATRIX

BY A. A. ALBERT

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CHICAGO

Communicated April 22, 1944

A correlation matrix is a symmetric matrix R with prescribed non-diagonal elements (correlation coefficients) and the matrix problem with which we are concerned is that of determining the diagonal elements (communalities) of R so that the resulting R will have minimum rank ρ. In a recent article[1] I defined the *ideal rank* r of R to be the largest order of a non-vanishing minor of R such that all elements of the minor are correlation coefficients. Evidently $\rho \geqq r$ and I gave a procedure for determining the communalities under the hypothesis that $\rho = r$.

Indeed we may always select our notation for the elements of M so that

$$R = \begin{pmatrix} A & B' & G' \\ B & C & H' \\ G & H & K \end{pmatrix},$$

where A, B, C are r-rowed square matrices and the matrix B has an inverse $B^{-1} = E$. Then I proved the

LEMMA 1: *If the ideal rank of R is r the matrix $K - GEH'$ is a diagonal matrix. The rank of R is r if and only if $K = GEH'$, $H = GEC$, $G = HE'A$, $B' = AEC$.*

In cases usually encountered in factor analysis the ideal rank r is much less than the order n of R and the r-columned matrices G and H normally have their *maximum* rank r. The purpose of the present note is to provide a proof of the rather astonishing

THEOREM. *Let the ideal rank r of R be also the rank of its submatrices G and H. Then there exist unique communalities such that the resulting correlation matrix R has rank r.*

We shall prove this result by several applications of the first statement in Lemma 1. Since K is symmetric it implies that GEH' is symmetric. Let the rank of G be r so that we may select a notation for R so that

$$G = \begin{pmatrix} S \\ P \end{pmatrix}, \quad H = \begin{pmatrix} T \\ Q \end{pmatrix}$$

for r-rowed square matrices S and T such that S is non-singular. Define a non-singular matrix

$$L = \begin{pmatrix} I & 0 \\ -PS^{-1} & I \end{pmatrix}$$

of $n - 2r$ rows, for appropriately selected identity matrices I, and form the products

$$LG = \begin{pmatrix} S \\ 0 \end{pmatrix}, \quad LH = \begin{pmatrix} T \\ W \end{pmatrix},$$

where $W = Q - S^{-1}PT$. Then

$$L(GEH')L' = (LG)E(LH)' = \begin{pmatrix} SE \\ 0 \end{pmatrix}(T'W') = \begin{pmatrix} SET' & SEW' \\ 0 & 0 \end{pmatrix}$$

is a symmetric matrix and necessarily $SEW' = 0$. Since SE is non-singular we have proved that $W' = 0$, $W = 0$,

$$LH = \begin{pmatrix} T \\ 0 \end{pmatrix}, L(GEH')L' = L\begin{pmatrix} SET' & SEQ' \\ PET' & PEQ' \end{pmatrix}L' = \begin{pmatrix} SET' & 0 \\ 0 & 0 \end{pmatrix}.$$

Suppose now that H has rank r so that LH has rank r and T must have rank r. Then T is non-singular and we may prove

LEMMA 2: *The matrices* $A - B'T^{-1}S$ *and* $C - BS^{-1}T$ *are diagonal matrices.*

For the submatrix of R consisting of its first $3r$ rows and columns is a correlation matrix

$$N = \begin{pmatrix} A & B' & S' \\ B & C & T' \\ S & T & U \end{pmatrix},$$

which is exactly like R except that K has been replaced by U, G by S, H by T. Every diagonal element of N is a diagonal element of M and B is an r-rowed non-zero minor of N. Then the ideal rank of N is r and, by the first part of Lemma 1, $U - SB^{-1}T'$ is a diagonal matrix. However, this result is actually a part of the statement that $K - GEH'$ is diagonal. We now obtain two correlation matrices

$$N_1 = \begin{pmatrix} U & T & S \\ T' & C & B \\ S' & B' & A \end{pmatrix}, \quad N_2 = \begin{pmatrix} A & S' & B' \\ S & U & T \\ B & T' & C \end{pmatrix}$$

of ideal rank r from N. The matrix N_1 is obtained by interchanging the first and third blocks of r rows and columns in N and the matrix N_2 by interchanging the second and third blocks of r rows and columns in N. Our lemma then follows from the first part of Lemma 1.

If R has rank r then N, N_1, N_2 have rank r. Applying Lemma 1 we have the uniqueness part of our theorem which we state as

LEMMA 3: *Let R have rank r. Then the communalities are uniquely determined by the equations*

$$A = B'T^{-1}S, \qquad C = BS^{-1}T, \qquad K = GB^{-1}H'. \qquad (1)$$

This lemma provides a modification of my earlier technique for computing the communalities. In the new technique we compute the inverses of the *three* matrices S, T, B and the diagonal elements of three matrix products.

To complete the proof of our theorem we assume now that R has ideal rank r and that r is the rank of G and H. By Lemmas 1 and 2 we may select the diagonal elements of R so that (1) is satisfied. Then

$$L(H - GEC) = \begin{pmatrix} T \\ 0 \end{pmatrix} - \begin{pmatrix} SEC \\ 0 \end{pmatrix} = \begin{pmatrix} T - SEC \\ 0 \end{pmatrix} = 0$$

since $C = BS^{-1}T$, $T = SEC$. Similarly

$$L(G - HE'A) = \begin{pmatrix} S \\ 0 \end{pmatrix} - \begin{pmatrix} TE'A \\ 0 \end{pmatrix} = \begin{pmatrix} S - TE'A \\ 0 \end{pmatrix} = 0$$

since $A = B'T^{-1}S$, $S = TE'A$. But L is non-singular and therefore $H = GEC$, $G = HE'A$. Finally, we have the identity

$$GE(B' - AEC)' = (H - GEC)E'A + G - HE'A$$

so that, since the right member is zero, $GE(B - AEC)' = 0$. The first r rows of this product form the matrix $SE(B' - AEC)' = 0$. Thus $(B' - AEC)' = 0$, $B' = AEC$, the rank of R is r. This completes the proof of our theorem.

Let us close by observing that R cannot have its ideal rank as its rank unless the ranks of G and of H are the same. For this is a consequence of the fact that when $H = GEC$ the rank of H is at most the rank of G, and that when $G = HE'A$ the rank of G is at most the rank of H. However, even when G and H have equal ranks *less than* r the minimum rank of R need not be its ideal rank r. Indeed let $G = H = 0$, K be a diagonal matrix and

$$B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad A = \begin{pmatrix} x_1 & 0 \\ 0 & x_2 \end{pmatrix}, \qquad C = \begin{pmatrix} y_1 & 1 \\ 1 & y_2 \end{pmatrix}.$$

Then every minor of R with more than two rows and containing no communality has a zero row and so $r = 2$. But $B' = AEC$ now becomes $C = A^{-1}$ and this is false since A^{-1} is a diagonal matrix and C is not a diagonal matrix.

---

[1] The matrices of factor analysis, *Proc. Nat. Acad. Sci.*, **30**, 90–95 (1944).

PROCEEDINGS
OF THE
NATIONAL ACADEMY OF SCIENCES
Volume 30 July 15, 1944 Number 7

Copyright 1944 by the National Academy of Sciences

X-RAY SENSITIVITY CHANGES IN MEIOTIC CHROMOSOMES AND THE NUCLEIC ACID CYCLE

By A. H. Sparrow*

Biological Laboratories, Harvard University

Communicated May 17, 1944

It is well known that the sensitivity of cells to x-ray injury is influenced by such factors as moisture content, the stage of nuclear or chromosomal development, the percentage of cells in mitosis and the degree of tissue differentiation. However, concerning the change of sensitivity during the nuclear cycle there is much disagreement. Practically all stages have been claimed by one or more authors to be the stage of maximum sensitivity. Resting stage or very early prophase has been claimed to be the most sensitive by Strangeways and Hopwood[1] and by Langendorf and Langendorf.[2] Prophase has been found most sensitive by Marshak,[3,4] Creighton[5] and Sax.[6] Carlson[7] found mid and late prophase most sensitive. Newcombe[8] states that sensitivity increases as the chromosomes approach metaphase while Sax and Swanson[9] have observed that sensitivity decreases during this period. Maximum injury to cells rayed at metaphase is reported by Mottram[10] and by Whiting[11] while Langendorf and Langendorf[2] and Marshak[12,13] claim metaphase and anaphase are least sensitive. Some authors have found two stages of high sensitivity, e.g., metaphase and anaphase by Holthusen,[14] Bozeman[15] and Sparrow,[16] anaphase and prophase by Regaud,[17] anaphase and telophase by Vintemberger[18] and telophase and prophase by Langendorf and Langendorf.[2]

From the results cited above it would appear that maximum sensitivity bears no constant relationship to the stage of mitosis at the time of raying. Yet the general aspects of mitosis and meiosis are remarkably uniform even in very distantly related forms. In the present paper meiotic metaphase and anaphase are shown to be stages of high sensitivity in *Trillium* and a plausible explanation for at least some of the disagreement with the results of other investigators is given. The relationship between nucleic acid concentration and chromosome radiosensitivity is also discussed.

Methods.—Rhizomes of *Trillium erectum* L. were obtained from commercial dealers during October and November. They were kept in a constant temperature room at 6–8°C. except when being examined or rayed. At this temperature meiosis proceeds slowly and a fairly accurate determination of stages is possible. First anaphase is probably the shortest stage but usually lasts at least one day. Meiotic prophase and microspore resting stages are very slow so that several weeks often elapsed between raying and smearing. Low temperature also seemed to reduce the amount of chromosome clumping and stickiness which often follow exposure to x-rays.

All plants were examined by removing part or all of one or more anthers for test smears. It was thus possible to select plants showing a high synchronization of division at the desired stage. This was further aided by the characteristic lack of a regular interkinetic resting nucleus between first and second divisions in this species.[19] The inhibiting effect of x-rays on spindle activity afforded an additional check on the stage rayed. Cells rayed at first metaphase usually developed into tetraploid microspores, those at first anaphase into diploid and those at second anaphase or later into the usual haploid form.

Cells rayed at meiotic prophase were scored for aberrations at first anaphase. Those rayed at metaphase and anaphase showed little or no increase in aberrations until microspore division which followed directly, without an intervening second meiotic division. Scoring was done at microspore metaphase following irradiation at meiotic metaphase and anaphase, and at microspore anaphase following x-ray treatment of microspores in the resting stage.

Results.—The data for doses of 50, 100 and 200 r at various stages of microsporogenesis are given in table 1. In most cases 50 or 100 cells were scored. In two cases (numbers 9 and 10) only 10 cells were counted but in these the very high number of fragments per cell compensates for the relatively small sample. The number of cells actually scored is indicated in table 1 under *n* but to facilitate comparisons all data have been expressed as the number of aberrations per hundred pollen mother cells.

Chromosomes rayed at first meiotic metaphase and anaphase appear practically normal when examined one, two or three days later. For example, in plant number 10 approximately 9 per cent of 300 first anaphases examined before raying showed a bridge or fragment or both, and when examined two days after exposure to 200 r less than 7 per cent of the cells showed aberrations. Obviously, therefore, the x-rays cause no immediate increase in the percentage of aberrations at first anaphase. However, there is a delayed reaction which becomes apparent in the microspore mitosis. For example, in plant number 9, which was given 200 r at first metaphase, a large number of giant (4*n*) pollen grains were later observed. In these so many fragments were present that it was generally difficult or impossible to tell which were centric and which were acentric fragments. The mean total of both counted at microspore metaphase was 65.2 ± 7.07. The corresponding mean for diploid microspores resulting from rayed first anaphases was 43.7 ± 3.45 or 87.4 per pollen mother cell. Deducting 20 (the normal microspore chromosome complement derived from each pollen mother cell) gives a corrected mean of 45.2 fragments for each pollen mother cell rayed at first metaphase and 67.4 for each one rayed at first anaphase. The mean number of fragments per pollen mother cell for the same dosage at leptotene-zygotene was 7.48 and 7.87 (plants 7 and 8). Obviously, therefore, the sensitivity of leptotene-zygotene is very much smaller than that of either metaphase or anaphase of the first meiotic division.

TABLE 1

THE NUMBERS OF BRIDGES, RINGS AND FRAGMENTS INDUCED BY RAYING AT DIFFERENT STAGES OF MICROSPOROGENESIS[a]

PLANT NUMBER	STAGE RAYED	DOSE IN r	n[a]	BRIDGES	RINGS	FRAGMENTS	TOTAL ABERRATIONS
1	Leptotene-zygotene	50	100	46	2	117	165
2	Pachytene	50	100	66	1	210	277
3	Leptotene-zygotene	100	50× 2	128	2	216	346
4	Late leptotene-zygotene	100	50× 2	200	4	580	784
5	Late pachytene	100	50× 2	340	6	556	902
6	Microspores (resting stage)	100	100× 4	100	8	492	600
7	Leptotene-zygotene	200	50× 2	288	16	748	1052
8	Leptotene-zygotene	200	100	230	12	787	1029
9	1st metaphase,	200	10× 10	..	[b]	4520	...
10	1st anaphase	200	10× 20	..	[b]	6740	...
11	Microspores (resting stage)	200	100× 4	256	12	776	1044

[a] Where fewer than 100 pollen mother cells (or the equivalent number of derived microspores) were scored, the data presented have been multiplied by the factor as given under n. Thus, the data can be compared directly.

[b] Rings were present but not recorded.

For stages rayed at different dosages the data cannot be compared directly. But since leptotene-zygotene stages were rayed at all dosages, the relative sensitivity at any other stage can be calculated with respect to leptotene-zygotene by assuming that all stages have similar dosage curves. Accordingly, leptotene-zygotene has been assigned an arbitrary unit sensitivity. Then, for any dosage, relative sensitivity is simply the number of aberrations at a given stage divided by the number of aberrations at leptotene-zygotene for the same dosage. The relative sensitivities given in table 2 are calculated from "total aberrations" in table 1 except for plants 9 and 10. In these calculations were made using fragments only, because scoring was done at microspore metaphase and hence bridges could not be included. The relative sensitivities show an increase from 1 at leptotene-zygotene to 1.7 at pachytene, 2.6 at late pachytene, 5.9 at first metaphase and 8.7 at first anaphase, and then drop again in the post-meiotic resting stage to approximately the same as early meiotic prophase.

TABLE 2

RELATIVE SENSITIVITY OF STAGES BASED ON DATA FROM TABLE 1

STAGE RAYED	RELATIVE SENSITIVITY
Leptotene-zygotene	1.0
Zygotene-early pachytene	2.3
Pachytene	1.7
Later pachytene	2.6
Metaphase I	5.9
Anaphase I	8.7
Microspore resting nucleus (plant no. 6)	1.7
Microspore resting nucleus (plant no. 11)	1.0

Discussion.—A number of recent papers point to the growing recognition of the importance of the nucleic acids in protein synthesis and nuclear metabolism. There is also evidence that small doses of x-rays can temporarily block one or more of the essential processes involved in the nucleic acid cycle.[20] This latter observation has prompted the following discussion of the significance of changes in relative nucleic acid concentrations in terms of chromosomal radiosensitivity.

Relatively small concentrations of desoxyribose nucleic acid are present in resting nuclei while metaphase and anaphase chromosomes contain very high concentrations.[21] Conversely, the high concentration of ribose nucleic acid present in cells undergoing, or about to undergo, rapid proliferation decreases[22, 23, 24] while the desoxyribose content increases.[25, 26] Specifically it has been shown that during meiotic prophase initial high concentrations of ribose nucleic acid fall off[24] and that the desoxyribose, which starts at a low level, increases as metaphase is approached.[26] Presumably the ribose nucleic acid, or ribose nucleotides, pass into the nucleus, are changed into the desoxyribose form, polymerized and incorporated into the structure of the chromosome, probably as nucleohistone.[27, 28] In support of this theory is the recent demonstration by Mitchell[20] that x- and gamma-radiations result in the accumulation of pentose nucleotides in the cytoplasm and inhibit the synthesis of desoxyribose nucleic acid. The inhibition is attributed to enzyme inactivation. Since the ribose nucleotides increase following irradiation and since they may be present in high concentrations at stages of low x-ray sensitivity and vice versa, it is clear that their concentration *per se* does not determine high radiosensitivity. For desoxyribose nucleic acid or nucleotides, however, the reverse is true. Stages, such as resting nuclei or early prophase, which have low concentrations of desoxyribose nucleotides or nucleic acid have a low sensitivity, while metaphase and anaphase, stages which have a high concentration, show a high radiosensitivity. This suggests a positive correlation between radiosensitivity and the concentration of desoxyribose nucleic acid or nucleotides. However, it does not necessarily mean that a high content *per se* determines high sensitivity. Other factors may determine high sensitivity and high nucleic acid content may be merely concomitant. This suggestion is in accord with Dale's[29] conclusion that the radiosensitivity of an enzyme can be modified by the amounts and radiosensitivities of other substances present.

Our results indicate that x-rays fail to break metaphase and anaphase chromosomes, as such, but that they somehow react to produce potential breaks which become visible chromatid or chromosome breaks only after the nucleus has passed through an interphase. It has been suggested that nucleic acid protects the metaphase chromosomes from x-ray injury[9, 12, 30] but this now seems improbable. However, the possibility remains that something inherent in the nucleoprotein structure of the chromosome temporarily prevents or inhibits the potential break from becoming a visible one. In this connection it may be significant that following irradiation at metaphase or anaphase breaks appear only after an interphase has occurred during which both the matrix and the nucleic acid structure undergo extensive alterations.

Experimental evidence concerning the physical-chemical changes induced in living cells by x-rays is meager but it is of interest that ultraviolet, which is known to produce qualitative genetic changes similar to those resulting from x-rays can cause a depolymerization of sodium thymonucleate.[31] Since breakdown of the nucleic acid macromolecule has also been suggested as a possible cause of ultra-violet induced mutations[31] it would be of considerable interest to determine whether or not irradiation at metaphase and anaphase would induce more mutations than the same dosage at prophase or resting stage.

Further evidence of the high sensitivity of metaphase is afforded by combined x-ray and colchicine treatments. Guyer and Claus[32] found cancer tissue in rats more sensitive to x-rays after injection of colchicine, and Brücke and Hueber[33] report that skin metastases treated with both colchicine and x-rays completely disappeared while metastases in the same patient not pretreated with colchicine showed a much less favorable response to irradiation. Further, Halberstaedter and Back[34] found that the lethal dose is 50 per cent smaller for Paramoecia which have been pretreated with colchicine. Since it is known that colchicine prolongs the metaphase by inhibiting spindle activity, it seems probable that the increased sensitivity to x-rays following colchicine treatment stems from an increase in the percentage of highly sensitive metaphases.

A relationship between malignancy and an abnormally high concentration of desoxyribose nucleic acid (hyperchromatism) has been noted by a number of authors.[35, 36, 37] In plants malignancy in certain types of pollen has also been attributed to abnormal nucleic acid metabolism associated with extra heterochromatic *B*-chromosomes.[38] Since the effectiveness of radiation therapy in cancer is dependent upon differential radiosensitivity it seems plausible that the high radiosensitivity of tumors may be associated with their high content of desoxyribose nucleic acid.

Support of the suggested relationship between desoxyribose nucleic acid and radiosensitivity is afforded by data on break frequency in heterochromatic regions which are characterized by differences in nucleic acid content.[39, 40, 41] White[42] found relatively few x-ray induced breaks in the negatively heterochromatic *X*-chromosome of spermatogonial divisions of *Locusta*, while several authors[43–46] report a significantly higher than random break frequency for positively heterochromatic regions in Drosophila. Since negative and positive heterochromatin are considered to be, respectively, under- and overcharged with desoxyribose nucleic acid with respect to the rest of the chromosomes[37, 41] this is further evidence that sensitivity to x-rays may be related to the concentration of desoxyribose nucleic acid.

From the data presented it seems clear that, in *Trillium*, meiotic metaphase and anaphase are stages of high sensitivity and early meiotic prophase a stage of low sensitivity to x-rays. The prophase to anaphase sensitivity change approaches a 1:9 ratio. Similar ratios have been reported for mitotic divisions in fertilized eggs. Mottram[10] found a 1:8 ratio between resting stage and metaphase in *Ascaris* ova, and Vintemberger[18] a 1:6 ratio between resting stage and telophase in frogs' eggs and a slightly lower value at anaphase. For *Tradescantia*, Sax[47] estimates that meiotic chromosomes are at least ten times as susceptible as those in microspore resting nuclei, but does not identify the meiotic stage of maximum sensitivity. However, metaphase and anaphase are not *generally* recognized as stages of high sensitivity even in recent papers dealing with relative sensitivity.[5, 30, 48–50] Yet evidence for high sensitivity of these stages is available from a remarkalby wide taxonomic range of organisms. These include Protozoa (Paramoecium),[34] Nemathelminthes (*Ascaris*),[10, 14] Insecta (*Sciara* and *Habrobracon*),[15, 11] Amphibia (frog),[18] Mammalia (rat),[32] and in plants the present work on *Trillium*. The observations include both mitotic and meiotic stages. It seems likely, therefore, that metaphase and anaphase stages may prove to be generally of higher sensitivity than prophase and resting stages.

Some of the earlier results are open to question because of the failure to examine the succeeding mitoses. In this connection it is probably significant that all of those who report metaphase or anaphase highly sensitive had examined material at some time after irradiation or used percentage of lethals as a measure of sensitivity. Many who report low sensitivity for these stages made cytological examinations but apparently failed to allow sufficient time between raying and fixation. Dividing cells observed while being irradiated were reported to complete mitosis without showing any adverse effects.[51] Here again it was probably a matter of not continuing observations sufficiently long, for the present investigation indicates that metaphase and anaphase chromosomes, as such, are highly resistant to x-ray breakage but that their high susceptibility to injury is abundantly evident if sought after in the next mitosis. It is therefore concluded that x-ray "hits" on metaphase and anaphase chromosomes do not induce immediate breaks but do induce potential breaks which are not actually realized until the following division. This delay in the appearance of the aberrations is probably responsible for at least some of the controversial claims regarding the radiosensitivity of metaphase and anaphase chromosomes.

Summary.—The relative effect of x-rays on different stages of meiosis and on microspore resting nuclei have been investigated in *Trillium.* Using the number of chromosome aberrations as a measure of sensitivity it was found that sensitivity is low in early meiotic prophase. Pachytene is roughly twice as sensitive, metaphase six and anaphase nine times as sensitive as early meiotic prophase. The sensitivity of post-meiotic resting nuclei is approximately that of early meiotic prophase. Since stages of high sensitivity to x-rays also have a high content of desoxyribose nucleic acid and stages of low sensitivity a low content, it appears that radiosensitivity can be correlated with nucleic acid metabolism. It is also suggested that non-random breaks in heterochromatic regions may be associated with a difference in the amount of desoxyribose nucleic acid in such regions as compared to the adjoining euchromatic chromatin. The relationship of tumor radiosensitivity to nucleic acid content is also briefly discussed.

Acknowledgment.—The author acknowledges financial assistance from The National Research Council and wishes to express his gratitude to Professor Karl Sax for his interest in the problem and reading of the manuscript.

---

[*] National Research Fellow 1942–1943.

[1] Strangeways, T. S. P., and Hopwood, F. L., *Proc. Roy. Soc. London,* **B100,** 283–293 (1926).

[2] Langendorf, H., and Langendorf, M., *Strahlentherapie,* **40,** 97–110 (1931).

[3] Marshak, A., *Jour. Gen. Physiol.,* **19,** 179–198 (1935).

[4] Marshak, A., these Proceedings, **25**, 510–516 (1939).

[5] Creighton, M., *Jour. Exp. Zool.*, **87**, 347–369 (1941).

[6] Sax, Karl, *Genetics*, **25**, 41–68 (1940).

[7] Carlson, J. G., *Jour. Morph.*, **71**, 449–462 (1942).

[8] Newcombe, H. B., *Jour. Genet.*, **43**, 145–171 (1942).

[9] Sax, Karl, and Swanson, C. P., *Amer. Jour. Bot.*, **28**, 52–59 (1941).

[10] Mottram, J. C., *Arch. Middlesex Hospt.*, **30**, 98–119 (1913).

[11] Whiting, A. R., *Jour. Exp. Zool.*, **83**, 249–269 (1940).

[12] Marshak, A., these Proceedings, **23**, 362–369 (1937).

[13] Marshak, A., *Proc. Soc. Exp. Biol. and Med.*, **38**, 705–713 (1938).

[14] Holthusen, H., *Pfluger's Arch.*, **187**, 1–24 (1921).

[15] Bozeman, M. L., *Genetics*, **28**, 71 (1943).

[16] Sparrow, A. H., *Rec. Genet. Soc. of Amer.*, **12**, 54–55 (1943).

[17] Regaud, C., *Bull. Assoc. franc. Cancer*, **12**, 482–487 (1923).

[18] Vintemberger, P., *Compt. Rend. Soc. Biol.*, **98**, 532–535 (1928).

[19] Sparrow, A. H., Huskins, C. L., and Wilson, G. B., *Can. Jour. Res.*, **C19**, 323–350 (1941).

[20] Mitchell, J. S., *Brit. Jour. Radiology*, **16**, 339–343 (1943).

[21] Caspersson, T., *Skand. Arch. Physiol.*, Suppl. No. 8, **73**, 1–151 (1936).

[22] Caspersson, T., and Thorell, B., *Chromosoma*, **2**, 132–154 (1941).

[23] Mirsky, A. E., in *Recent Advances in Enzymology*, Interscience Publishers, Inc., New York, Vol. 3, pp. 1–34, (1943).

[24] Painter, T. S., *Bot. Gaz.*, **105**, 58–68 (1943).

[25] Brachet, J., *Arch. Biol.*, **44**, 519–576 (1933).

[26] Caspersson, T., *Chromosoma*, **1**, 147–156 (1939).

[27] White, M. J. D., in *Cytology and Cell Physiology*, Oxford University Press, Oxford pp. 139–159, (1942).

[28] Mirsky, A. E., and Pollister, A. W., *Biol. Symposia*, **10**, 247–260 (1943).

[29] Dale, W. M., *Biochem. Jour.*, **36**, 80–85 (1942).

[30] Swanson, C. P., *Jour. Gen. Physiol.*, **26**, 485–494 (1943).

[31] Hollander, A., Greenstein, J. P., and Jenrette, W. V., *Jour. Nat. Cancer Inst.*, **2**, 23–28 (1941).

[32] Guyer, M. F., and Claus, P. E., *Proc. Soc. Exp. Biol. and Med.*, **42**, 565–568 (1939).

[33] Brücke, F. th. v., and Hueber, E. F. v., *Klin. Wchnschr.*, **18**, 1160–1161 (1939).

[34] Halberstaedter, L., and Back, A., *Nature*, **152**, 275–276 (1943).

[35] Biesele, J. J., Poyner, H., and Painter, T. S., *University of Texas Publication*, no. 4243, pp. 1–68 (1942).

[36] Stowell, R. E., *Jour. Nat. Cancer Inst.*, **3**, 111–121 (1942).

[37] Koller, P. C., *Nature*, **151**, 244–246 (1943).

[38] Darlington, C. D., and Thomas, P. T., *Proc. Roy. Soc. London*, **B130**, 127–150 (1941).

[39] Schultz, J., and Caspersson, T., *Arch. f. Exper. Zellforsch.*, **22**, 650–654 (1938).

[40] White, M. J. D., *Jour. Genet.*, **40**, 67–82 (1940).

[41] Darlington, C. D., *Nature*, **149**, 66–69 (1942).

[42] White, M. J. D., *Proc. Roy. Soc. London*, **B119**, 61–84 (1935).

[43] Prokofyeva-Belgovskaya, A. A., *C.R. (Doklady) Acad. Sci. U.R.S.S.*, **23**, 270–272 (1939).

[44] Kaufmann, B. P., these Proceedings, **25**, 571–577 (1939).

[45] Hinton, T., *Genetics*, **27**, 119–127 (1941).

[46] Muller, H. J., *Cold Spring Harbor Symp. on Quant. Biol.*, **9**, 151–167 (1942).

[47] Sax, Karl, *Genetics*, **23**, 494–516 (1938).

[48] Marshak, A., *Proc. Soc. Exp. Biol. and Med.*, **39**, 194–198 (1938).

[49] Froier, K., Gelin, O., and Gustafsson, A., *Bot. Notiser*, **1941**, 199–216.

[50] Koller, P. C., *Proc. Roy. Soc. Edinburgh*, **B61**, 398–429 (1943).

[51] Canti, R. G., and Donaldson, M., *Proc. Roy. Soc. London*, **B100**, 413–419 (1926).

THE SPORE DISCHARGE MECHANISM OF COMMON FERNS

By Allen L. King

Department of Physics, Dartmouth College

Communicated June 6, 1944

Among ferns the Polypodiaceae are most common and are highly developed. In this family the mature sporangium, rather uniformly, is a flattened obovoid capsule supported on a long pedicel (*P*, Fig. 1). The unicellular thin side walls enclose the region in which spores develop. Along approximately two-thirds of the edge of this lenticular form, a single row of nearly cuboid cells (*A*, Fig. 1) become differentiated. This incomplete ring, the annulus, is prominent on the back and apex of the sporangium and is attached to the pedicel at the rear. From the annulus to the front of the pedicel a number of thin-walled cells complete the ring. Two of these cells meet in a straight line, the stomium (*S*, Fig. 1), and when the sporangium dehisces, cleavage first takes place along this juncture.

The annulus consists of thirteen or more cells with thickened inner and radial walls and thin flexible outer and side walls. The inner walls form a continuous band with crenations due to a slight arching inward at each cell. This band is very elastic and acts as a spring to expel the spores with considerable speed at the time of dehiscence.

When the sporangium becomes mature, evaporation of water from the annular cells induces sufficient stress to cause cleavage at the stomium. The line of cleavage follows cell boundaries toward the rear of the sporangium. Frequently weak lateral cells are split in two.[1] Meanwhile, the annulus slowly straightens and continues to evert sometimes to the extent that the two ends of the recurved annulus almost meet. Nearly all the spores are carried on the free end in a cup formed by the lateral walls of the sporangium. Suddenly the annulus snaps back to its original position and in the process it may catapult spores to distances of 1 to 2 cm.[2]

Although the function of the annulus in spore discharge was known prior to 1700, the details of the mechanism were not understood until the end of the last century. It was recognized that adhesion of the cell contents to the walls within annular cells is greater than cohesion. Careful observations revealed that the appearance of a bubble within the cell initiated the sudden snap of the annulus; and it was suspected that rupture of the cell sap released the elastic energy of the bent annulus. In 1915 Renner[3] and Urpsrung[4] independently measured the negative pressure required for bubble formation by two methods. They obtained values of 200 to 300 atmospheres.

Size of Sporangia.—The annular cell consists of thick base and radial walls and thin flexible side and outer walls. During dehiscence the base becomes recurved, the radial walls remain relatively unchanged in shape although they become reoriented, and the weak side and outer walls are pulled in. For the following analysis the cross-section of an annular cell is assumed to have the form shown in figure 2. Since the outer wall of length m is observed to be almost flat, the initial pressure inside the cell probably differs from that outside by not more than a few atmospheres. The base of length a has a radius of curvature r' and the radial walls have a height b. Just before rupture the base becomes recurved until it has a radius r. Membrane m, which is relatively inextensible, takes on a semicylindrical form provided end effects are neglected. Width l of the cell is assumed to be unaltered by the process.

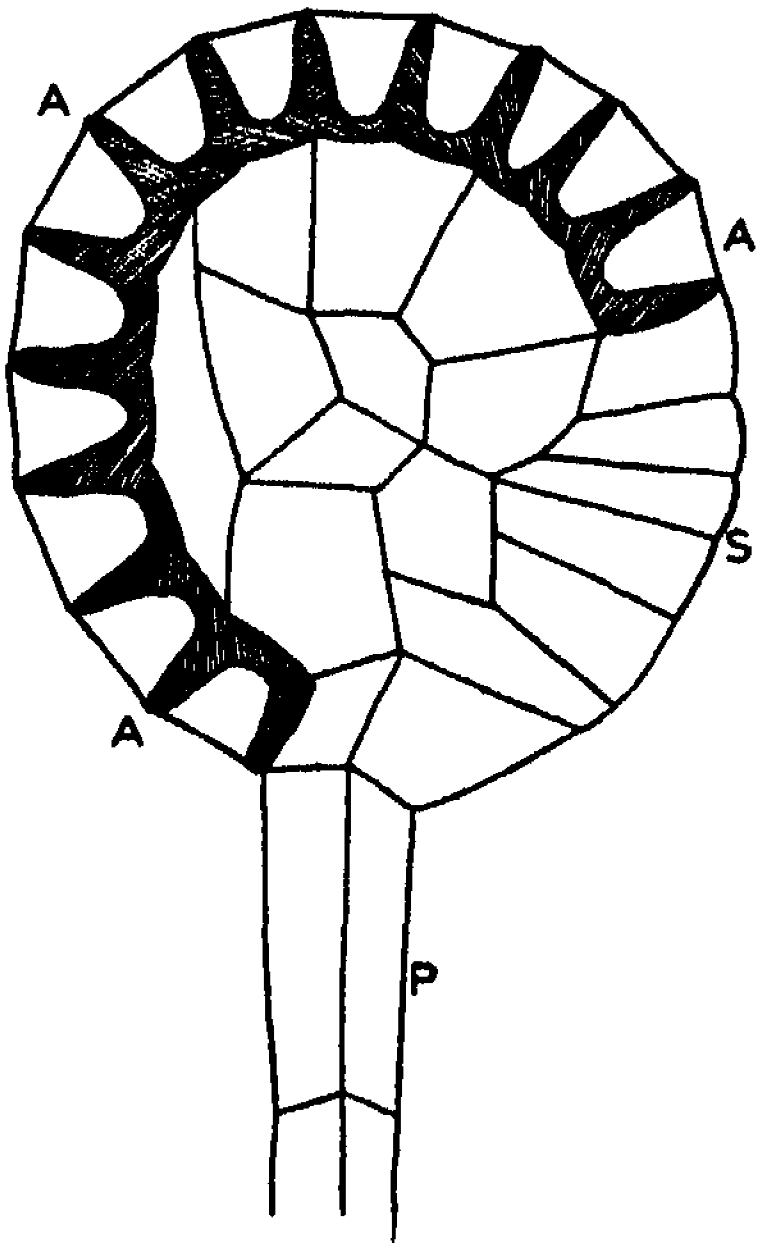


FIGURE 1

Side view of a sporangium showing location of the annulus A, the stalk or pedicel P and the stomium S. The hatched portions represent the thickened walls of the annular cells. Although spores usually may be seen through the transparent side walls of a mature sporangium, they have been omitted in this drawing.

Suppose the difference in pressure across membrane QR is q, directed normal to the surface as shown in figure 3. A tension T, which is constant over the entire membrane, will be set up along QR. Each element of area $l\,ds$ of the membrane is in equilibrium so that

$$T\,d\phi = q\,ds \qquad (1)$$

Since T and q are constants, Equation (1) defines a circle of radius T/q. Thus the outer wall of the annular cell becomes a cylindrical surface of length l and radius T/q. Since the membrane is semicircular in cross-section at the instant of rupture, this radius equals m/π so that the tension T has the value qm/π and makes an angle $\alpha = a/2r$ with the radial wall.

From figure 2, r' is found to be given by

$$m = 2(r' + b) \sin (a/2r') \tag{2}$$

and r by

$$m = \pi(r - b) \sin (a/2r). \tag{3}$$

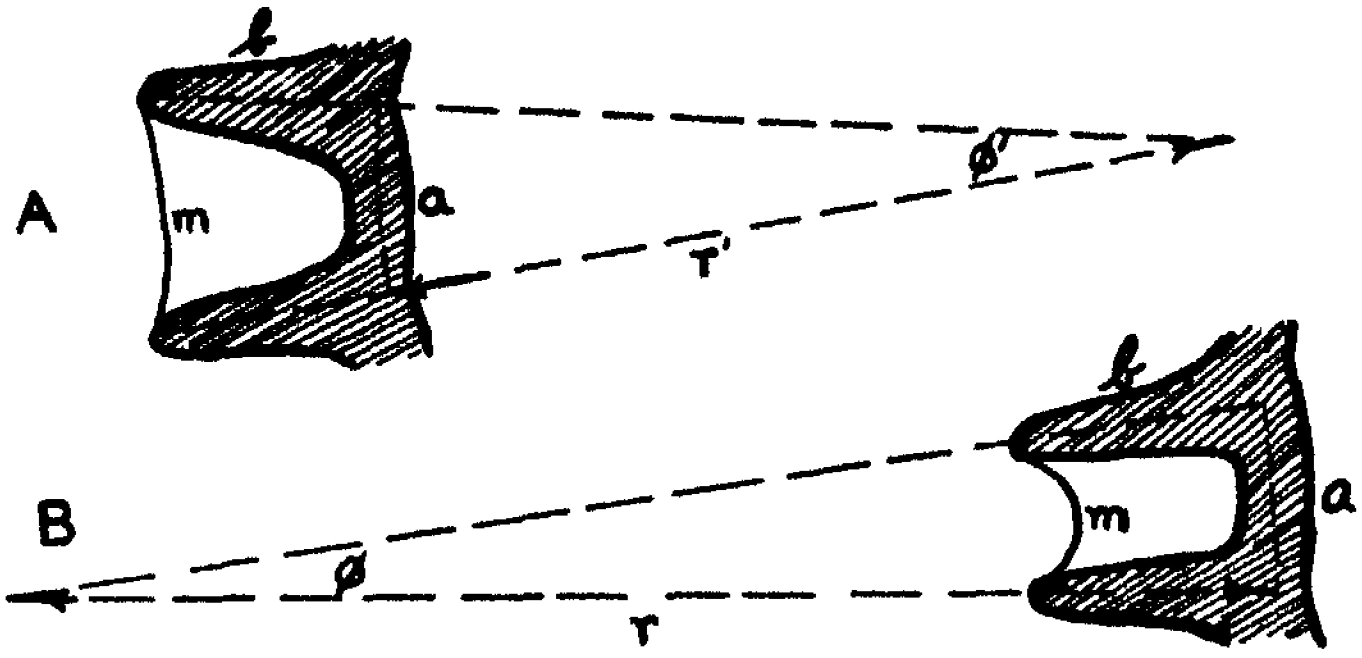


FIGURE 2

Cross-section of a single annular cell before dehiscence, *A*; at the moment of full eversion, *B*. *m* is a thin flexible membrane through which water is believed to pass during dehiscence.

The first terms in the expansions of the sine functions are sufficient for this analysis. On eliminating m from the resulting relations,

$$\frac{1}{r} = \frac{0.364}{b} - \frac{0.636}{r'}. \tag{4}$$

Values of r/b, r'/b and r'/r are entered in table 1. The product rr' has a minimum value when $r = 5.49b$ and $r' = 3.49b$.

TABLE 1

RELATION BETWEEN INITIAL AND FINAL RADII OF SPORANGIUM

r'/b	r/b	r'/r	r'/b	r/b	r'/r
1.75	∞	0.00	5.0	4.2	1.2
2.0	22.2	0.09	6.0	3.9	1.5
3.0	6.6	0.46	7.0	3.7	1.9
4.0	4.9	0.82	8.0	3.5	2.3

Since the outer wall becomes semicylindrical before bubble formation, the annulus must straighten out or recurve. The minimum value of r'/b, therefore, is 1.75 corresponding to $r = \infty$. If θ' and θ represent the initial and final angles subtended by *all* the annular cells, then

$$r'\theta' = r\theta. \tag{5}$$

The maximum value of θ is 2π, since the annulus can recurve only until its two ends meet. On the other hand, θ' usually is not less than π. For these limiting values of θ and θ' the ratio $r'/r = 2$. In table 1 this ratio is seen to occur when r'/b is slightly greater than seven. Thus the radius of a ripe sporangium is not less than 1.75, nor greater than 7 times the height of an annular cell. As shown in table 2, average values of r'/b for seven species of ferns lie well within this range.

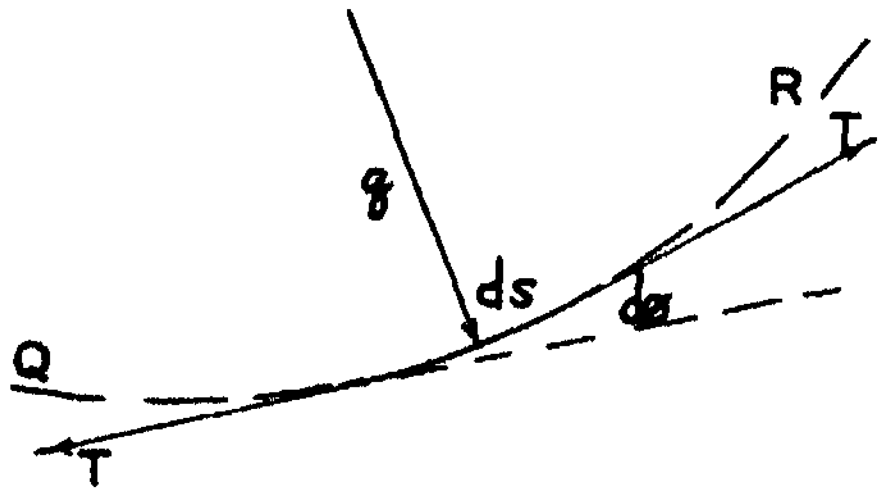


FIGURE 3

Cross-section of a portion of membrane m at the moment of full eversion showing element of length ds. The vector labeled q represents the force $ql\,ds$ and the vectors labeled T represent the force $T\,dl$.

TABLE 2

SIZES OF SPORES, SPORANGIA AND ANNULAR CELLS

SPECIES	b	a/b	c/b	r'/b*	l/b	e/b†
Cyrtomium falcatum	30 μ	0.63	0.27	3.5	2.0	1.3
Dryopteris uliginosa	31 μ	0.74	0.26	2.6	2.1	1.1
Phlebodium aureum	44 μ	0.69	0.18	3.8	1.8	1.0
Pteris cretica	38 μ	0.59	0.22	3.8	1.5	1.4
Pteris longifolia	47 μ	0.58	0.25	4.8	1.4	1.0
Nephrolepis pectinata	28 μ	0.68	0.29	4.0	1.6	0.9
Nephrolepis exaltata	29 μ	0.76	0.25	2.6	2.0	1.0
Averages	35 μ	0.67	0.25	3.6	1.8	1.1

* Usually sporangia have elliptical cross-sections. Major and minor axes were measured and r' was computed for the middle of the annulus.

† Frequently, spores are slightly ellipsoidal. Only the non-fertile spores of *Nephrolepis exaltata* (Boston fern) showed considerable variation in size.

Young's Modulus of the Annulus.—Consider a single annular cell $OPQR$ of a fully everted sporangium, as shown in figure 4. Membrane QR is a cylinder of radius T/q, as shown above, where q is the negative pressure of the cell sap on the membrane (the pressure outside is negligible). q acts uniformly over the entire inner surface of the cell. For the membrane QR, however, it may be replaced by the tension T acting along the tangent to the membrane at the juncture with the radial walls.

The moment of the forces on cell surfaces OPQ about O consists of the moment of T, the moment of q over surfaces OP and PQ and the bending moment of the elastic forces in the base OP. The effect of the weight of cells attached to PQ is very small compared with the other stresses and may be neglected. To the approximations involved here the bending moment of the base equals $EI(r + r')/rr'$, where E is Young's modulus of the base material, I is the moment of area for a transverse section of the base and r and r' are the radii of curvature before and at the instant of full eversion. The annulus is assumed to be in equilibrium just before rupture occurs, so that

$$EI(r + r')/rr' = ql\{\tfrac{1}{2}b^2 + r(r - b)(1 - \cos\phi)\} + Tl\{r \sin(\alpha - \phi) - (r - b)\sin\alpha\}. \quad (6)$$

If the transverse section of the cell base has height c, the moment of area I about the central line is $lc^3/12$. With this value of I and for small angles, equation (6) reduces to

$$E = 6qb^2rr'/c^3(r + r'). \quad (7)$$

With $c = b/4$, $r = 5.5b$, $r' = 3.5b$, and $q = 200$ atmospheres, Young's modulus for the annulus comes out approximately 1.6×10^{11} dynes/cm.^{2} This value is about 25 per cent greater than that reported by Press[5] for viscose rayon yarn. Common green woods[6] have moduli ranging from 1.0 to 1.5×10^{11} dynes/cm.^{2} Since the thickened wall of the annular cell is especially spring-like, its modulus of elasticity should be somewhat greater than that for common wood fibers. The ratio of Young's modulus to the density of the annular walls is greater than the same ratio for many other solids.

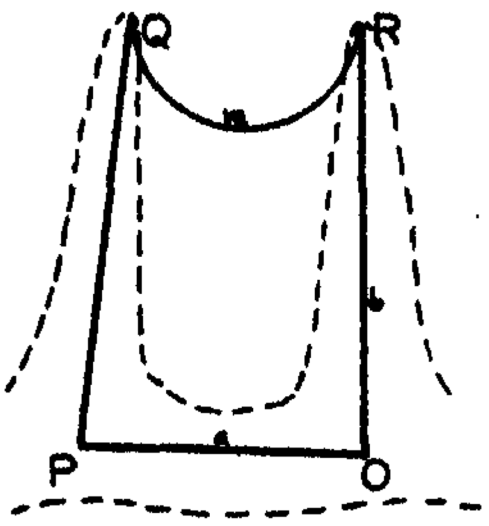


FIGURE 4

Cross-section of a single annular cell at full eversion. For the analysis in this paper the cell is assumed to have the form $OPQR$. The dotted lines indicate the actual outline of the cell walls.

Initial Speed and Range of Spores.—Since the thickened base of the annulus is a continuous structure, it may be considered a circular arc fastened to the pedicel at one end. When the sporangium is fully everted sixty odd spores are held within a cup formed by the lateral walls which remain fixed to the free end of the annulus. The total potential energy of the everted annulus is

$$W = \tfrac{1}{2}naEI(1/r + 1/r')^2, \quad (8)$$

where n represents the number of cells in the annulus and the other symbols have already been defined.

When the vapor bubble forms, the negative pressure is assumed to be entirely released; so that the annulus springs back with a snap to its original position. It is further assumed that the cup with spores is fastened to the end cell of the annulus. Even if the cup is three or four cells from the end, the initial speed and range of the spores are increased by less than 100 per cent. As the cup slips from behind it, the spore mass acquires momentum directed away from the sporangium. To a first approximation, the radius of the annular arc at the moment the spores leave is

$$R = rn/(n - 1). \tag{9}$$

This relation arises if the angle subtended by the annular arc from pedicel to spore mass, initially $(n - 1)a/r$ and finally na/R, is supposed to remain unchanged during the discharge process. The decrease in potential energy experienced by the annulus as the radius changes from r to R is

$$V = (aEI/r^2)(1 + r/r' - 1/2n). \tag{10}$$

In order to obtain an upper limit for the average speed v of spore ejection, this change in potential energy is assumed to be transformed entirely into kinetic energy of the spore mass. The kinetic energy of the moving annulus and cup and all energy losses are neglected. The kinetic energy gained by 60 spores is

$$U = 5\pi\rho e^3v^2, \tag{11}$$

where ρ is the density and e is the diameter of each spore.

Equate U and V, eliminate EI, and solve for v. Because $r'/2n(r + r')$ is much less than unity, the average speed comes out

$$v = (aqlb^2/10\pi\rho e^3r)^{1/2}. \tag{12}$$

On substituting average values of a, b and e from table 2, $r = 5.5b$, $\rho = 1$ g./cc. and $q = 200$ atmospheres, the speed of ejection of the spores is calculated to be 10 meters/sec.

By application of Stoke's law, the distance a spore may travel is given by

$$d = ve^2(\rho - \rho')/18\eta, \tag{13}$$

where v is the initial speed, e is the diameter and ρ the density of the spore, and η is the viscosity and ρ' the density of air.[7] $\eta = 0.00018$ poise at 20°C., $\rho = 1$ g./cc., and ρ' may be assumed negligible. Using these data and the average value of e from table 2 (nearly 40 microns) in equation (13), the range of spores is computed to be approximately 5 cm. Although this result is more than twice the observed values, the discrepancy is not unexpected. Energy losses were neglected, the mass of the spore cup was not taken into account and the problem was otherwise much idealized.

If the negative pressure q were less than 10 atmospheres or greater than

10,000 atmospheres, Young's modulus of the annular arc and the range of spores would have come out too small or too large. A sporangium can evert again and again by alternate imbibition of water and deshiscence after the spores have been expelled, even in the case of herbarium specimens more than fifteen years old.[8] This suggests that the annular cells of a spore-filled mature sporangium contain water and that this water is ruptured at the time of bubble formation. *The above computations, therefore, yield additional evidence for the belief that the tensile strength of water is around 200 atmospheres.* Haller[9] obtained 200–300 atmospheres by means of a piezo-electric device in a stream of water where cavitation occurred. Measurements by Dixon[10] and by Vincent[11] using Berthelot's method gave values as high as 150–200 atmospheres. Many experimental and theoretical values are in marked disagreement with these results; but an extensive discussion and evaluation of them is beyond the scope of this paper.

The author here expresses his appreciation for the use of the greenhouse and the helpful advice given by Professor Charles J. Lyon of the botany department at Dartmouth College. Also he is much indebted to Drs. L. J. Briggs and L. B. Tuckerman of the National Bureau of Standards for pointing out two errors in the original manuscript.

[1] Atkinson, G. F., *The Study of the Biology of Ferns by the Collodion Method*, Macmillan and Co., New York (1894), p. 68.

[2] Ingold, C. T., *Spore Discharge in Land Plants*, Oxford University Press, London (1939), p. 119.

[3] Renner, O., *Jahr. f. Wiss. Botanik*, **56**, 647 (1915).

[4] Ursprung, A., *Ber. d. Deutsche Bot. Gesell.*, **33**, 153 (1915).

[5] Press, J. J., *Jour. App. Phys.*, **14**, 224 (1943).

[6] *International Critical Tables*, Vol. II (1927), p. 6.

[7] Buller, A. H. R., *Researches on Fungi*, Vol. 1, Longmans, Green, and Co. (1909), p. 190.

[8] Atkinson, G. F., op. cit., p. 71.

[9] Haller, P. de, *Schweizerische Bauzeitung*, **101**, 260 (1933).

[10] Dixon, H. H., *Transpiration and the Ascent of Sap in Plants*, Macmillan and Co., London (1914).

[11] Vincent, R. S., *Phys. Soc. Proc.*, **53**, 126 (1941); Vincent, R. S., and Simmonds, G. H., *Phys. Soc. Proc.*, **55**, 376 (1943).

SCALE CURVES IN CONFORMAL MAPS[1]

By Edward Kasner and John De Cicco

Departments of Mathematics, Columbia University, Illinois Institute of Technology

Communicated May 26, 1944

1. Let a surface Σ be mapped in point-to-point fashion upon a plane π with cartesian coördinates (x, y). The scale function $\sigma = ds/dS$, which is the ratio of the corresponding elements of arc length in π and in Σ, depends, in general, not only upon the point but also upon the direction. It is independent of the direction if and only if the mapping is conformal.

We wish to present some theorems concerning the scale function σ in any general conformal mapping of a surface Σ upon a plane π, and to apply the results to the cartography of the sphere. The classic projections of Ptolemy, Mercator and Lambert appear in new light. A *scale curve* on Σ or π is the locus of a point along which the scale function σ does not vary. Therefore in the conformal case, there are ∞^1 scale curves, defined by the finite equation $\sigma(x, y) = \text{const.}$ We shall be chiefly concerned with these scale curves. In the non-conformal case, there is always a double infinity of scale curves. This will be studied in our next paper.

2. Kasner has proved that the complete system of ∞^2 isogonal trajectories of a given base family (M) of ∞^1 curves upon a surface Σ is *linear* (in the analytic sense, as distinguished from the algebraic sense) if and only if (M) is an isothermal family.[2]

This theorem is important in cartography for the following reasons. Suppose it is desired to traverse a surface Σ always cutting a given base family (M) upon Σ at a constant angle (thus defining *general loxodromes* with respect to (M) on Σ) in such a way that when Σ is mapped upon a plane π, the path described is represented in π by a straight line (for any angle). The above result states that this can be accomplished if and only if (M) is an isothermal family of curves.

If, in addition, it be desired that the base family (M) consist wholly of geodesics upon Σ, then Σ must be applicable to a surface of revolution and the family (M) must correspond to meridians. If Σ contains more than one such isothermal system (M) of geodesics, then it must contain ∞^2 such systems (M) and the surface Σ must be of constant gaussian curvature G.[3]

3. Any family of ∞^1 curves can represent the scale curves of a conformal map upon a plane π of any one of a certain class of surfaces Σ. In the present paper, we shall say that a family of curves is of the type λ if it represents the scale curves of a conformal map of a surface Σ upon a plane π such that the gaussian curvature G is constant along each of the scale curves. A family of curves $f(x, y) = \text{const.}$, is of the type λ if and only if f satisfies a certain partial differential equation of the fourth order. In that

event, it is possible to find a function $\lambda = \lambda(f)$ such that λ satisfies the partial differential equation: $\lambda_{xx} + \lambda_{yy} = \phi(\lambda)$. The logarithm of the scale function σ also satisfies a differential equation of the same form.

4. Any isothermal family is of the type λ but not every family of the type λ is isothermal. Thus the family of similar ellipses $x^2 + 2y^2 =$ const. is of the type λ but is not isothermal. We may use the term *quasi-isothermal* for the general λ type.

If the ∞^1 scale curves of a conformal map of a surface Σ upon a plane π form an isothermal system such that the gaussian curvature G is constant along each of the scale curves, then Σ is either developable or applicable upon a non-developable surface of revolution. In the former case, the scale curves may be any isothermal family (and the conformal map is of perfectly general character), but in the latter case the scale curves must be either parallel straight lines or concentric circles (and the conformal map is of special character).

5. A family of the type λ (quasi-isothermal) is parallel if and only if it consists of parallel straight lines or concentric circles. The associated surface Σ must then be either developable or applicable upon a surface of revolution.

If a developable surface Σ is mapped conformally upon a plane π with parallel curve scales, then the map is the product of an unrolling of Σ upon π by one of the transformations $Z = z^n$, $Z = \log z$, $Z = e^z$, followed by a similitude. This is entirely different from the isothermal case where the conformal map is perfectly arbitrary.

If a surface Σ of constant non-vanishing gaussian curvature is mapped conformally upon a plane π with isothermal or parallel scales, then the scales must be parallel straight lines or concentric circles. *Thus the only conformal maps of a sphere upon a plane with isothermal or parallel scales are the Mercator projection (with straight scales) and the Ptolemy stereographic and Lambert projections (with circular scales).*

6. A famous theorem of Beltrami states that if a surface Σ is mapped upon a plane π in some point-to-point fashion such that its ∞^2 geodesics are represented by straight lines in π, then Σ is of constant gaussian curvature. We shall study now the situation where not all the geodesics correspond to straight lines, but the maximum number of lines arise.

Kasner has proved that under any arbitrary point-to-point map of a surface Σ upon a plane π (or more generally, upon another surface Σ^1), there are at most $3\ \infty^1$ geodesics which correspond to straight lines (or geodesics on Σ^1). This is called partial geodesic representation.

In a conformal map, there are at most ∞^1 geodesics on Σ which correspond to straight lines on π (since the $2\ \infty^1$ minimal lines on Σ already correspond to the $2\ \infty^1$ minimal lines on π). If a conformal map carries exactly ∞^1 geodesics of Σ into straight lines of π, *then the scale curves of the* *conformal mapping must form a parallel curve family*, and the orthogonal family is the required system of geodesics.[4]

If a surface Σ of constant gaussian curvature is conformally mapped upon a plane π such that ∞^1 geodesics of Σ are pictured by straight lines of π, then the scale curves are parallel lines or concentric circles. For the cases of a developable surface or a sphere, the maps are the ones given in Section 5.

7. The only families of ∞^1 lines or circles which are quasi-isothermal are the pencils of straight lines or circles. The associated surface Σ must be either developable or applicable to a surface of revolution. This is a large generalization of a theorem of Lagrange.

If a developable surface Σ is mapped conformally upon a plane π such that the scales are straight (or circular), the map is the product of an unrolling of Σ upon π by the transformation $Z = z^n$ (or either $Z = \log z$ or $Z = e^z$) followed by a similitude.

The only conformal map of a sphere upon a plane with straight scales is the Mercator projection. The only conformal maps of a sphere upon a plane with circular scales are the Ptolemy stereographic and Lambert projections.

8. The sphere cannot be mapped conformally on a plane so that the geodesics become straight. Geodesic circles can be mapped into circles, but not so that concentric families remain concentric (this is true even for non-conformality).

The quasi-isothermal type of family of curves (for both surface and plane) includes the Laplace, Poisson and Helmholtz-Pockels equations arising in the theory of heat conduction.[5]

---

[1] Presented to the American Mathematical Society, February, 1944. Some of our cartographic results have been outlined (for both the conformal and non-conformal maps) in *Science* **98**, 324–325 (1943), and in *Science News Letter*, March 25, 1944.

[2] For a recent unified discussion of this theorem, see the paper by De Cicco, "New Proofs of the Theorems of Beltrami and Kasner on Linear Families," *Bull. Am. Math. Soc.*, **49**, 407–412 (1943). The original proof appeared in *Math. Annalen*, 1904.

[3] Kasner, "Isothermal Systems of Geodesics," *Trans. Am. Math. Soc.*, **5**, 55–60 (1904).

[4] Kasner, "The Problem of Partial Geodesic Representation," *Ibid.*, **7**, 200–206 (1906).

[5] See Kasner, "Geometry of the Heat Equation," *Proc. Nat. Acad. Sci.*, **19**, 257 (1933). See. p. 261 for Fialkow's fundamental geometric property of fourth order, characterizing the quasi-isothermal type.

THE SPAN OF THE TRANSLATIONS OF A FUNCTION IN A LEBESGUE SPACE*

BY I. E. SEGAL

BALLISTIC RESEARCH LABORATORY, ABERDEEN PROVING GROUND, MD.

Communicated May 16, 1944

The refined harmonic analysis of a real or complex-valued function on the real line[1] is connected with the determination of the span,[2] in a particular space, of the translations of the function. That this span depends mainly on the zeros of the Fourier transform of the function was first indicated by Wiener.[3] He proved that, for the spaces L and L_2, the span is the entire space if and only if the Fourier transform of the function is, in some sense, non-vanishing. At the same time he raised the question of how far the analogue to this is true in a general Lebesgue space,[4] L_p. The present note is concerned with this question for the type of non-vanishing that seemed appropriate for the Fourier transform of an element of such a space, i.e., non-vanishing except on a set of Lebesgue measure zero. We show that non-vanishing of the Fourier transform is sufficient for spanning when $p > 2$, but insufficient when $1 < p < 2$. The precise result is stated in the following theorem.[5]

Let f be an element of L_p whose L_p-Fourier transform exists, where $1 < p < \infty$; and let (A) and (B) be conditions on f, as follows:

(A) The real zeros of the L_p-Fourier transform of f constitute a set of Lebesgue measure zero;

(B) The translations of f span L_p.

If $p \geq 2$, (A) implies (B). If $1 < p < 2$, (A) does not imply (B)—there exist elements of L_p for which (A) but not (B) holds; however, (B) implies (A).

For succinctness of statement, some matters of definition have been deferred; these will now be taken up.

1. The Lebesgue space L_p is either that of real or that of complex-valued functions, on the real line, in its strong topology.

2. A translation of an element, $f(\cdot)$, of L_p is a function of the form $f(x + a)$, where a is a constant.

3. In case $p > 2$ the L_p-Fourier transform of an element, f, of L_p is said to exist if f is the L_q-Fourier transform of an element of L_q, where[6] $p^{-1} + q^{-1} = 1$. (N. B.—This last relation implies $1 < q < 2$; also, for $1 < p < 2$ the L_p-Fourier transform always exists.)

Instead of giving a full proof, which would require excessive space, a condensed version is presented.[7]

LEMMA 1: *An element of L_p, $f(\cdot)$, satisfies (B) if and only if the following is true: if $g(\cdot) \in L_q$ and $\int_{-\infty}^{\infty} f(x + y)\overline{g(y)}dy = 0$ for all real values of x, then $g(\cdot) = 0$.*

By a result in the theory of Banach spaces,[8] the translations of $f(\cdot)$ span L_p if and only if no non-zero continuous linear functional vanishes on all translations. By Riesz' representation theorem,[9] every continuous linear functional on L_p, λ, has the form: $\lambda f(\cdot) = \int_{-\infty}^{\infty} f(x)\overline{g(x)}dx$, where $g(\cdot)$ is a particular element of L_q. The proof is completed by combining these facts in straightforward fashion.

At this point it is logical to divide the proof into two parts, according as $p \geq 2$ or $1 < p < 2$; the respective proofs, as well as the results, differ in general character.

Case 1: $p > 2$. This case is distinctly the simpler. The proof is indirect; suppose that $f(\cdot)$, a function satisfying the hypothesis, fulfills (A) but not (B). By Lemma 1 there exists a non-zero element of L_q, $g(\cdot)$, such that $\int_{-\infty}^{\infty} f(x + y)\overline{g(y)}dy = 0$ for all real values of x. Since $1 < q \leq 2$, the L_q-Fourier transform of $g(\cdot)$, $G(\cdot)$, exists; by hypothesis, the L_p-Fourier transform of $f(\cdot)$, $F(\cdot)$, exists. Hence the Parseval formula for Fourier transforms[10] applies, in the following manner: $\int_{-\infty}^{\infty} f(x + y)\overline{g(y)}dy = \int_{-\infty}^{\infty} e^{-ixy}F(y)\overline{G(y)}dy$. Noting that $F(y)\overline{G(y)}$ is an element of L, it follows from uniqueness theorems for Fourier transforms, and the fact that $f(\cdot)$ fulfills (A), that $g(\cdot) = 0$, a contradiction. This completes the proof, in Case 1.

Case 2: $1 < p < 2$. The proof that (A) does not imply (B) is based on two lemmas.

LEMMA 2: *Let N be a closed set of real numbers. Then there exists a function, f, that is both in L and in L_2, and whose L-Fourier transform vanishes at each point of N, and at no other real point.*

In case N is the complement of a finite open interval, there exist functions satisfying the conclusion of the lemma, and, moreover, real and non-negative; this fact follows trivially from a formula proved by Wiener.[11] For the general case, put $N = \bigcap_{k=1}^{\infty} N_k$, where N_k is the complement of some finite open interval. Let f_k be a function both in L and L_2 whose L-Fourier transform is real and non-negative and vanishes exactly on N_k. Now set $= \sum_{k=1}^{\infty} e_k f_k$, where e_k is chosen positive and so that the series is convergent both in L and L_2; e.g., take $e_k = k^{-2}||f_k||^{-1}$, where $||h(\cdot)|| = \int_{-\infty}^{\infty} |h(x)|dx + (\int_{-\infty}^{\infty} |h(x)|^2dx)^{1/2}$, for an arbitrary measurable function $h(\cdot)$. It is plain that ff fulfills the condition of the lemma.

LEMMA 3: *Let q satisfy $q > 2$. Then there exists a function of bounded* *variation, $K(\cdot)$, that: (1) is purely singular and continuous, (2) has spectrum,[12] N, of Lebesgue measure zero, (3) whose Fourier-Stieltjes transform, $g(\cdot)$, is an element of L_q.*

This lemma follows from a theorem of Salem's,[13] which states: if e is a positive number, then there exists a function of bounded variation on the interval $(0, 2\pi)$, $K(\cdot)$, that is purely singular, continuous, with spectrum of measure zero, and such that $g(n) = 0(n^{-1/2+e})$, where $g(t) = \int\limits_0^{2\pi} e^{itx} dK(x)$. It will be shown that this function $g(\cdot)$ is an element of L_q for every $q > \frac{2}{1-2e}$; since e can be arbitrarily small, the proof will then be complete.

Defining $n (= n(t))$ by $n - {}^1/_2 < t < n + {}^1/_2$, and integrating by parts, $g(t) = e^{2\pi i(t-n)} g(n) - i(t-n) \int_0^{2\pi} e^{i(t-n)x} (\int_0^x e^{inu} dK(u)) dx$. Taking absolute values, and applying Minkowski's inequality, $(\int\limits_{-\infty}^{\infty} |g(t)|^q dt)^{1/q} \leq (\sum\limits_{n=-\infty}^{\infty} |g(n)|^q)^{1/q} + (\int_{-1/2}^{1/2} t^q dt)^{1/q} (\sum\limits_{n=-\infty}^{\infty} (h(n))^q)^{1/q}$, where $h(n) = \int_0^{2\pi} |\int_0^x e^{inu} dK(u)| dx$. Therefore, it is sufficient, now, to show that $\{g(n)\}$ and $\{h(n)\}$ are both in l_q, where q is an arbitrary number greater than $\frac{2}{1-2e}$ (a general element of l_q is, here, a doubly infinite sequence $s(n)$ $(n = 0, \pm 1, \ldots)$ for which $\sum\limits_{n=-\infty}^{\infty} |s(n)|^q$ is convergent).

That $\{g(n)\} \epsilon\, l_q$ is plain from Salem's theorem. To show $\{h(n)\} \epsilon\, l_q$, a formal analogue to Parseval's theorem is employed, obtaining:

$$\int_0^x e^{inu} dK(u) = \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} (\int_0^x e^{-imu} du)(\int_0^{2\pi} e^{i(m+n)u} dK(u));$$

this equation will now be established rigorously. Clearly, it is sufficient to show the following: if $S(\cdot)$ is a continuous function of bounded variation on $(0, 2\pi)$ such that $\{s(n)\} \epsilon\, l_t$ for some $t < \infty$, where $s(n) = \int_0^{2\pi} e^{inu} dS(u)$, then $\int_0^x dS(u) = \frac{1}{2\pi} \sum\limits_{m=-\infty}^{\infty} s(m)(\int_0^x e^{-imu} du)$. Let $f(x, r) = \sum\limits_{n=-\infty}^{\infty} r^{|n|} e^{inx}$, where $0 < r < 1$. It is easy that:

$$(1)\ f(x, r) = \frac{1 - r^2}{1 - 2r \cos x + r^2};\quad (2)\ S(x, r) = \int_0^{2\pi} f(x - y, r) dS(x)$$

is absolutely continuous, for r constant; (3) the Fourier-Stieltjes coefficients of $S(x, r)$ are $\{r^{|n|} s(n)\}$; (4) $\frac{\partial}{\partial x} S(x, r) \epsilon\, L_2(0, 2\pi)$, for r constant; (5) $\int_0^x d_u S(u, r) = \frac{1}{2\pi} \sum\limits_{m=-\infty}^{\infty} (\int_0^x e^{-imu} du) s(m) r^{|m|}$; (6) $|(\int_0^x e^{-imu} du) s(m) r^{|m|}| \leq d(m)|s(m)|$, where $d(0) = 2\pi$ and $d(m) = 2m^{-1}$ for $m \neq 0$; (7) $\sum\limits_{m=-\infty}^{\infty} d(m)|s(m)| < \infty$; (8) $\sum\limits_{m=-\infty}^{\infty} (\int_0^x e^{-imu} du) s(m) r^{|m|} \to \sum\limits_{m=-\infty}^{\infty} (\int_0^x e^{-imu} du) s(m)$, as $r \to 1$;

(9) $\int_0^x f(u - v, r)du$ is bounded for $0 < x < 2\pi$, $0 < v < 2\pi$, $0 < r < 1$, and converges to 1 or 0 as $r \rightarrow 1$ according as $x > v$ or $x < v$, respectively; (10) $\int_0^x d_u S(u, r) \rightarrow \int_0^x dS(u)$, as $r \rightarrow 1$. Combining (5), (8) and (10) yields the stated equation.

The proof that $\{h(n)\} \epsilon l_q$ can now be completed. Taking absolute values in the expression for $\int_0^x e^{inu}dK(u)$ and integrating shows that $h(n) \leq \sum_{m=-\infty}^{\infty} d(m)|g(m + n)|$ ($d(\cdot)$ as defined in (6), preceding paragraph). It is plain that $\{g(n)\} \epsilon l_u$ for every $u > \frac{2}{1 - 2e}$, and $\{d(n)\} \epsilon l_v$ for every $v > 1$. Applying a theorem of Weil's[14] concerning the spaces that contain the convolution of two functions in given spaces shows that $\{h(n)\} \epsilon l_q$.

Now the proof that (A) does not imply (B), in Case 2, will be concluded.[15] Let $g(\cdot)$, $K(\cdot)$ and N be as in Lemma 3. By Lemma 2 there exists a function, $f(\cdot)$ that is an element of both L and L_2, and whose L-Fourier transform, $F(\cdot)$, vanishes exactly on N. It is trivial that $f(\cdot)$ is an element of L_p and satisfies (A); that $f(\cdot)$ fails to satisfy (B) will now be shown. By the Fubini theorem, $\int_{-\infty}^{\infty} f(x + y)\overline{g(y)}dy = \int_{-\infty}^{\infty} e^{-ixy}F(y)d\overline{K(y)}$. Since $F(\cdot)$ vanishes on the spectrum of $K(\cdot)$, the latter integral is zero. From Lemma 1, it follows that (B) is not valid.

To prove that (B) implies (A) in Case 2, suppose this false; let $f(\cdot)$ be an element of L_p that satisfies (B), but not (A). From the latter fact it follows that there exists a measurable set of finite positive measure, S, on which the L_p-Fourier transform of $f(\cdot)$, $F(\cdot)$, vanishes. Let $G(\cdot)$ be the characteristic function of S. Plainly $G(\cdot) \epsilon L_p$; therefore $G(\cdot)$ is the L_q-Fourier transform of an element of L_q, $g(\cdot)$. Using the Parseval formula as in the proof for Case 1, it follows that $\int_{-\infty}^{\infty} f(x + y)\overline{g(y)}dy = 0$ for all real values of x. Therefore, by Lemma 1, $f(\cdot)$ does not satisfy (B), a contradiction.[16]

---

[*] The work reported in the present note, except for the present proof of Lemma 3, was done from May to July of 1940, largely at Yale University.

[1] A complete analogue to our results holds for functions on an arbitrary locally compact Abelian group. The proof of these general results is formally the same as in the case of the real line. A full statement and proof will be given in a paper to be offered to the *Trans. Amer. Math. Soc.*

[2] By the span of a set of element we mean the least closed linear manifold containing the set.

[3] Wiener, N., "Tauberian Theorems," *Ann. Math.*, s. 2, **33**, 1–100 (1932).

[4] *Ibid.*, p. 93. Wiener stated that he suspected an analogue was probably true for $1 < p < 2$ and possibly true more generally.

[5] This theorem and its proof can both readily be extended to the case of the translations of a set of functions; the case of a single function seems fundamental, however, and we therefore postpone treating the case of a set of functions to the paper referred to in footnote 2.

[6] This relation defines q for the rest of the paper, also.

[7] A more detailed proof will be included in the paper referred to in footnote 2.

[8] Banach, S., *Théorie des opérations linéaires*, Warsaw, 1932, p. 57; "Lemme."

[9] Riesz, F., "Untersuchungen über Systeme integrierbarer Funktionen," *Math. Ann.*, **69**, 449–497 (1910); p. 475.

[10] See Titchmarsh, E. C., *Theory of Fourier Integrals*, Oxford (1937), p. 105.

[11] Wiener, N., *The Fourier Integral and Certain of Its Applications*, Cambridge, England (1933), pp. 49–50.

[12] The spectrum of a function of bounded variation consists of all points such that the total variation of the function is positive over every neighborhood of the point.

[13] Salem, R., "On Singular Monotonic Functions of the Cantor Type," *Jour. Math. and Phys.*, **21**, 69–82 (1942).

[14] Weil, A., "L'Integration dans les groupes topologiques et ses applications," *Act. Sc. et Ind.*, No. 869, Paris, p. 55 (1940).

[15] We are pleased to thank J. D. Tamarkin and A. Zygmund for suggesting a modification in the statement of Lemma 2 which results in a shortening of this part of our proof.

[16] It seems noteworthy that this argument holds for $p = 2$ also, and together with the proof for Case 1, provides a new and somewhat simpler proof of Wiener's theorem for the space L^2.

ON THE SOLUTION OF A GENERAL TRANSFORM

BY I. S. REED

DEPARTMENT OF MATHEMATICS, CALIFORNIA INSTITUTE OF TECHNOLOGY

Communicated May 15, 1944

The purpose of this paper is to give a brief extension to the solution of a Watson transform with an unsymmetric kernel. The solutions with symmetrical kernels are already well known from the work that has been done by Hardy, Watson, Titchmarsh, Goodspeed and others.

THEOREM. *If,*

(*i*) *$\alpha(u)$ and $\beta(u)$ are reciprocals of integral functions which take real values on the real axis.*

(*ii*) *$r(u)$ is an integral function such that if $u = v + iw$*

$$\left|\frac{r(u)}{\alpha(1+u)}\right| < C_1 e^{Av+B|w|}, \quad \left|\frac{r(-u-1)}{\beta(1+u)}\right| < C_2 e^{Cv+D|w|}$$

$$0 < B < \pi, \ 0 < D < \pi.$$

(*iii*) *$k(s) = \alpha(s)/\beta(1-s)$, $h(s) = \beta(s)/\alpha(1-s)$ satisfy $k(s)h(1-s) = 1$ and are the Mellin transforms of $K(x)$, $H(x)$, respectively.*

(*iv*) *$\dfrac{k(^1/_2 + it)}{^1/_2 - it}$, and $\dfrac{h(^1/_2 + it)}{^1/_2 - it}$ belong to $L^2(-\infty, \infty)$.*

Then $\sum_{n=0}^{\infty}\frac{(-1)^n r(n)x^n}{\alpha(n+1)}$ *converges for* $0 < x < e^{-A}$ *and is represented by* $f(x)$ *of* $L^2(0, \infty)$ *for all positive* x, *and* $\int_0^{\infty} f(x)K(xy)dx = g(y)$ *where* $g(y)$ *is of* $L^2(0, \infty)$ *and is the analytic continuation for all positive* y *of* $\sum_{n=0}^{\infty}\frac{(-1)^n r(-n-1)^n}{\beta(n+1)^n}$ *which converges for* $0 < y < e^{-C}$. *Also* $\int_0^{\infty} g(x)H(xy)dx = f(y)$ *for* $y > 0$. *Hence* $f(x)$ *and* $g(x)$ *are the general transforms of the unsymmetrical kernels* $K(x)$ *and* $H(x)$.

Proof. Now by Cauchy's theorem,

$$f(x) = \sum_{n=0}^{\infty} (-1)^n \frac{r(n)x^n}{\alpha(n+1)} = (^1/_2\pi i)\int_{a-i\infty}^{a+i\infty} \frac{\pi}{\sin \pi u}\,\frac{r(-u)}{\alpha(1-u)}x^{-u}du$$

when $0 < a < 1$. By condition (iv) the integral representation holds first for $0 < x < e^{-A}$ and by analytical continuation it converges uniformly in any interval $0 < x_0 \leq x \leq X$.[1] Also $f(x) = 0(x^{-a})$ for all positive x. Thus by moving the line of integration across the origin, changing the integral only by a constant, and then taking $a < -{}^1/_2$ for the lower limit of integration and $a > -{}^1/_2$ for the upper limit, we see that $\int_0^{\infty}\{f(x)\}^2dx < \infty$ or is integrable in the Lebesgue sense $L^2(0, \infty)$.

It follows then by Theorem 129 of Titchmarsh's *Fourier Integral* and its extension to unsymmetrical formulae mentioned in Section 8.9 that from condition (iv) and from $f(x)$ being a function of $L^2(0, \infty)$, that the formula $g(y) = \int_0^{\infty}f(x)K(xy)dx$ defines almost everywhere a function $g(y)$ belonging to $L^2(0, \infty)$ and that also the reciprocal formula $f(y) = \int_0^{\infty}g(x)H(xy)dx$ holds.

In particular,

$$\begin{aligned}\int_0^{\infty} f(x)K(xy)dx &= (^1/_2\pi i)\int_{a-i\infty}^{a+i\infty}\frac{\pi du}{\sin(\pi u)}\,\frac{r(-u)}{\alpha(1-u)}\int_0^{\infty}K(xu)x^{-u}dx\\ &= (^1/_2\pi i)\int_{a-i\infty}^{a+i\infty}\frac{\pi du}{\sin(\pi u)}\,\frac{r(-u)}{\alpha(1-u)}\,k(1-u)y^{u-1}\\ &= (^1/_2\pi i)\int_{a-i\infty}^{a+i\infty}\frac{\pi du}{\sin(\pi u)}\,\frac{r(-u)}{\beta(u)}\,y^{u-1}du\\ &= -(^1/_2\pi i)\int_{a-1-i\infty}^{a-1+i\infty}\frac{\pi dv}{\sin(\pi v)}\,\frac{r(-v-1)}{\beta(v+1)}\,y^v dv\\ &= \sum_{n=0}^{\infty}(-y)^n\frac{r(-n-1)}{\beta(n+1)} = g(y).\end{aligned}$$

Thus the proof is complete.

Example. When $-{}^1/_2 \leq \nu < {}^1/_2$ it is well known that if $K(x) = x^{1/2} Y\nu(x)$ then

$$H(x) = x^{1/2} H\nu(x)^2$$

and in this case if

$$\alpha(s) = 2^{s/2}\Gamma({}^1/_2\nu + {}^1/_2 s + {}^1/_4)\sec({}^1/_2\nu - {}^1/_2 s + {}^3/_4)\pi$$

then

$$\beta(s) = 2^{s/2}\Gamma({}^1/_2\nu + {}^1/_2 s + {}^1/_4)\csc({}^1/_2 s + {}^1/_2\nu + {}^1/_4)\pi$$

hence if $N = {}^1/_2 n + {}^1/_2\nu + {}^3/_4$

$$f(x) = \sum_{n=0}^{\infty} (-x/\sqrt{2})^n r(n) \cos\,({}^1/_2\nu - {}^1/_2 n + {}^1/_4)\pi/\Gamma(N)$$

$$g(x) = \sum_{n=0}^{\infty} (-x/\sqrt{2})^n r(-n-1) \sin\,({}^1/_2\nu + {}^1/_2 n + {}^3/_4)\pi/\Gamma(N).$$

If, moreover, $r(n) = \sec({}^1/_2\nu - {}^1/_2 n + {}^1/_4)$, then $r(-n-1) = \sec({}^1/_2\nu + {}^1/_2 n + {}^3/_4)\pi$ and so

$$g(x) = \sum_{n=0}^{\infty}(-x/\sqrt{2})^n \tan\,(N\pi)/\Gamma(N) = \lambda(x) \tan\,({}^1/_2\nu + {}^3/_4)\pi + \mu(x) \cot\,({}^1/_2\nu + {}^3/_4)\pi$$

where

$$\lambda(x) = \sum({}^1/_2 x^2)^n/\Gamma({}^1/_2\nu + n + {}^3/_4), \quad \mu(x) = \sum_{n=0}^{\infty} ({}^1/_2 x^2)^{n+1/2}/\Gamma({}^1/_2\nu + n + {}^5/_4).$$

Also

$$f(x) = \sum_{n=0}^{\infty} (-x/\sqrt{2})^n/\Gamma(N) = \lambda(x) - \mu(x).$$

Again, if

$$r(-n-1) = \csc\,(N\pi), \text{ then } r(n) = \csc\,({}^1/_2\nu - {}^1/_2 n + {}^1/_4)\pi$$

and

$$g(x) = \lambda(x) - \mu(x), f(x) = \lambda(x) \cot\,({}^1/_2\nu + {}^1/_4)\pi + \mu(x) \tan\,({}^1/_2\nu + {}^1/_4)\pi.$$

Also it is known that[3]

$$\lambda(y) - \mu(y) = \int_0^{\infty} \{\lambda(x) - \mu(x)\}\, (xy)^{1/2} J\nu(xy) dx$$

so we obtain the identity

$$\begin{aligned}\lambda(y) - \mu(y) &= \int_0^\infty \{\lambda(x) - \mu(x)\} (xy)^{1/2} J\nu(xy)dx \\ &= -\int_0^\infty [\lambda(x) \cot (1/2\nu + 1/4)\pi + \mu(x) \tan (1/2\nu + 1/4)\pi] \\ &\qquad (xy)^{1/2} H\nu(xy)dx \\ &= \int_0^\infty [\lambda(x) \cot (1/2\nu + 1/4)\pi + \mu(x) \tan (1/2\nu + 1/4)\pi] \\ &\qquad (xy)^{1/2} Y\nu(xy)dx\end{aligned}$$

also

$$\int_0^\infty [\lambda(x) \cot (1/2\nu + 1/4)\pi + \mu(x) \tan (1/2\nu + 1/4)\pi] (xy)^{1/2}[H\nu(xy) + Y\nu(xy)]dx = 0.$$

---

[1] Hardy, G. H., "Ramanujan and the Theory of Fourier Transforms," *Quart. Jour. Math.* (Oxford series), 8, 245–254 (1937).

[2] Titchmarsh, E. C., *Theory of Fourier Integrals*, p. 215 (1937), Oxford.

[3] Hardy, G. H., *Ramanujan*, p. 208.

PROCEEDINGS
OF THE
NATIONAL ACADEMY OF SCIENCES

Volume 30 August 15, 1944 Number 8

Copyright 1944 by the National Academy of Sciences

A SPECIFIC ABNORMALITY ASSOCIATED WITH A VARIETY OF GENOTYPES

BY L. C. DUNN AND S. GLUECKSOHN-SCHOENHEIMER

DEPARTMENT OF ZOOLOGY, COLUMBIA UNIVERSITY*

Communicated June 13, 1944

We have found recently a hitherto undescribed hereditary syndrome of abnormalities in the house mouse, in which a similar set of developmental changes in the gut and urogenital system appears as the expression of any one of about a dozen different genotypes, comprising seven different mutations with marked effects on the axial skeleton and tail. The results of a preliminary study suggest that the abnormality arises whenever some process in early development is sufficiently deflected from its normal path and that the different mutations affect this process in different degrees. It can be inferred from the genetical study that the process involves an inductive relationship by which the normal development of the posterior gut and urogenital system is connected with the development of the axial system.

An account of the morphology of the syndrome and of a detailed analysis of its inheritance and expression in some of the genotypes is being published elsewhere. The present note records the occurrence of the abnormality as a common type of response to a number of mutant changes.

The syndrome was first noted in one family of tailless mice in which certain offspring which were alive at birth lacked both anal and genital apertures and died shortly after birth. Further study of such abnormal animals from this and other stocks showed that the essential feature of this abnormality in all stocks is the simultaneous failure of development of the posterior parts of the gut and of the urinary duct, often resulting in an abnormal condition in which colon and urethra remain connected, forming a cloaca which, however, has no external opening. Nearly always the rectum is absent or vestigial and the anus fails to form. In some cases, kidneys, ureters, genital ducts and bladder are missing or abnormal but these are not constant parts of the basic syndrome and the frequency of these abnormalities varies in different genotypes. The syndrome is diagnosed at birth by the imperforate anus, absence or underdevelopment of the genital papilla and complete failure of urinary excretion. It is probably the latter feature which brings about the early death of animals with this syndrome. This abnormality has been found regularly in the following genotypes at birth:

1. *Sd Sd*. *Sd* is a dominant mutation producing a short tail and kidney defects in heterozygotes; homozygotes *Sd Sd* always show the full syndrome of imperforate anus and true cloaca without excretory duct and in addition are tailless and generally have no kidneys (cf. Gluecksohn-Schoenheimer[1]).

2. Tt^0 *uu*. Tt^0 is tailless (cf. Dunn[2] for review of effects of *T* and related tail mutations); *u* is a new recessive which in combination with Tt^0 produces the imperforate anus and urogenital syndrome in all or nearly all *uu* animals.

T+ *uu*—short-tailed animals which show the syndrome in a large proportion but not in all homozygotes.

$^{++}$ *uu*—animals with short contorted tails, weak, sterile and stunted; but showing the syndrome in a small proportion (about 10 per cent) of homozygotes.

3. *Ki T*/++; *Ki* (Kink tail) is a mutation closely linked with *T*, and producing in combination with *T* a shortened kinky tail or complete taillessness (Dunn and Caspari[3]). The syndrome has appeared in about half of the animals of this genotype, often accompanying an abnormality of the spinal column, spina bifida aperta, in the sacral region. It appears also in animals of genotype *Ki*/*T* but its frequency in this genotype has not been established.

4. *Ki T*/+ t^x; this genotype is always tailless, frequently shows spina bifida aperta, and more than half of the newborn animals have the imperforate syndrome and die shortly after birth.

5. *Ki T*/++; *Sd*/+; tailless, often with syndrome and with absence or abnormality of kidneys more severe than is found in *Sd*/+.

6. *Ki*/+; *Sd*/+; tailless; occasionally like 5.

7. *Fu*/*Fu*; *Fu* is a tail mutation closely resembling *Ki*. It acts as a semilethal, some *Fu Fu* being viable. *Fu* is closely linked with *Ki* and *T* (Dunn and Caspari[3]). In nearly all *Fu Fu* animals at birth (25 out of 26 dissected) the kidneys are missing or reduced; while 9 out of 26 showed the imperforate syndrome.

8. *Fu*/t^0 × *Fu*/t^0. The offspring of this cross are *Fu*/*Fu* or *Fu*/t^0 since t^0 is lethal and prevents recombination in this area (Dunn and Caspari, in press). Many of the offspring show the same conditions as described in 7 above.

9. *Fu*/+; *Sd*/+. This combination frequently shows the imperforate syndrome and nearly always has kidney defects more severe than

those associated with *Sd* alone. It appears that *Fu* and *Sd* show some cumulative effect in increasing the extent of urogenital abnormality.

10–11. In litters containing the genotypes $\frac{Fu\ +}{+\ T}$ and $\frac{Ki\ T\ +}{+\ +\ Fu}$, animals with the syndrome and suspected to be of the above genotypes have been found. Because such animals die shortly after birth it has been impossible to verify the diagnosis of genotype.

In general, it can be seen that the same syndrome or parts of it tend to appear when *u*, *Sd*, *Ki* and *Fu*, are present in combination with each other or with T, t^0 or t^1. It does not appear when only the latter mutations are present. The frequency with which the syndrome appears in the different genotypes is variable and probably depends in part on factors, other than those named, in which the animals differ, since it has not yet been possible to prepare isogenic stocks. For example, when the *Ki T* chromosome was derived from one male about half of the offspring which received it showed the syndrome, while of those receiving it from other parents, a much smaller proportion were abnormal. In spite of this variability, it is evident that animals homozygous for *Sd Sd* always show the full syndrome, while those with *uu* Tt^0 nearly always show it. One other homozygous combination, *Fu Fu*, is sometimes viable until after birth, and although it shows the imperforate condition in only about a third of the newborns, all of them show related effects on the kidneys. Homozygosity for one of the three factors *Sd*, *u* or *Fu* thus produces a high probability that urogenital development will be abnormal. The other mutations involved, *Ki T*, t^0 and t^1, are all lethal before birth when homozygous and their effects on producing the syndrome can only be observed in heterozygous combination. It is noteworthy that of the many simple heterozygotes observed, *Ki*/+, *T*/+, t^0/+, t^1/+, *Fu*/+, *u*/+ and *Sd*/+ none has been found with this abnormality. On the other hand, it is often found in animals heterozygous for *two or more* of these mutations, *e.g.*, *Ki T*; *Ki T*/t^0 or t^1; *Fu*/t^0 or t^1; *Fu*/*T*; *Fu*/+*Sd*/+; *Ki*/+*Sd*/+, etc., indicating that one dose of each of these is able to increase the effect of one dose of any other one, in a manner similar to the interaction of two doses of the same allele in the viable homozygotes such as *Sd Sd*. The frequency of abnormal animals, however, appears to be lower among heterozygous combinations than among homozygotes.

These facts may be explained on the following hypothesis. Each of the mutations concerned has a quantitative effect on some process concerned with development of the posterior spine tending to cause it to fall below its normal level. This produces abnormalities or absence of tail vertebrae in all combinations of these mutations, and abnormalities in the sacral region at least in most combinations. (Spina bifida aperta is regularly found in *Sd Sd*, often in *Ki T*/t^0 or t^1, *Ki*/*T*, *Ki T*/*Fu*, *Fu*/t^0 or t^1, *Fu/Fu.*) When the posterior axial region is sufficiently defective, it fails to provide the normal inductive stimuli which regulate the development of the posterior gut and urinary ducts, and particularly the separation of the urogenital sinus from the rectum. This assumption is based not on embryological evidence but on the evident connection between abnormalities of the posterior spine and urogenital system.

The hypothesis implies that the mutations studied (*u*, *Sd*, *Ki*, *T*, *Fu*, t^0 or t^1) have some effect in common which underlies the process or processes leading to the abnormal connection between gut and urinary duct. They apparently differ in their effects on certain other processes since each mutant probably has peculiar effects on tail form by which it can be identified. Further study may, of course, reveal differences in the urogenital effects of the different mutants, but the essential features of the syndrome, failure of posterior gut and urinary duct and absence of anus, appear to be the usual type of response of the embryo to anatomical defects brought about by a number of independent mutations.

* The experiments reported were carried out with the aid of a grant from the Josiah Macy, Jr., Foundation of which grateful acknowledgment is made.

[1] Glueckson-Schoenheimer, S., *Genetics*, **28**, 341–348 (1943).

[2] Dunn, L. C., *Growth*, **5**, 141–161 (1941).

[3] Dunn, L. C., and Caspari, E., these PROCEEDINGS, **28**, 205–210 (1942).

[4] Reed, S. C., *Genetics*, **22**, 1–13 (1937).

---

*CAN SPECIFIC MUTATIONS BE INDUCED BY SEROLOGICAL METHODS?**

BY A. H. STURTEVANT

THE W. G. KERCKHOFF LABORATORIES OF THE BIOLOGICAL SCIENCES, CALIFORNIA INSTITUTE OF TECHNOLOGY, PASADENA, CALIF.

Communicated July 5, 1944

It has been suggested by Haldane (1935) and by Irwin (1939) that many of the erythrocyte antigens of birds and mammals are rather direct gene products. The reason for this view is that (if we except a portion of the antigens present in species-hybrids of doves) there is a one-to-one correspondence between the presence of specific single genes and specific antigens (see summaries for man by Boyd, 1939, and Wiener, 1935, for rabbits by Castle and Keeler, 1933, for fowls by Todd, 1935, for doves and pigeons by Irwin, 1939). In these cases there is—with the exception noted—no evidence for gene interaction, either between allelomorphs or between genes at different loci.

The same conclusions may be applied to the genes and antigens responsible for the specificity of transplantation compatibility in mammals. The results of Loeb and Wright (1927) with guinea pigs show that transplantation is successful only if all the specific genes of the donor are present in the host. This evidently means that the host forms antibodies to any antigen present in the donor tissue, unless the same antigen is present in the host also. The genetic results mean further that these antigens are numerous, and that here also there is no gene interaction. Similar conclusions may be drawn from the experiments on tumor transplantation in mice.

As pointed out by Haldane (loc. cit.), this conclusion suggests that there may be some common configuration in the molecular structure of an antigen and that of the gene responsible for it. If this common configuration concerns that part of the antigen that is responsible for its specificity—and such a relation is implied in the suggestion—there are certain further possibilities to be investigated.

As is well known (see discussion by Landsteiner, 1936), there are two distinct properties of an antigen—that of inducing the formation of a specific antibody, and that of reacting with the specific antibody. These two properties evidently depend on the same specific configuration; but for the first it is usually (probably always) necessary that the specific configuration be present in or chemically attached to a protein, while for the reaction with the antibody such an association is not necessary.

These considerations lead to the supposition that, if a particular gene is responsible for the formation of a given antigen, then there is a possibility that the antibodies induced by this antigen may react with the gene. If this possibility exists there is a series of consequences that are of interest to the geneticist.

One such consequence is that it becomes possible to interpret the old experiments of Guyer and Smith (1920, 1924) with rabbit lenses. As is well known, the lens protein has specific antigenic properties. Guyer and Smith reported that they immunized fowls to rabbit lenses, and injected the immune sera into rabbits. There resulted lens defects in the offspring, and these defects were inherited, both through females and through males. These experiments have been criticized, and Finlay (1924), Huxley and Carr-Saunders (1924) and Ibsen and Bushnell (1934) have failed in attempts to confirm them. I am, however, informed by Professor R. R. Hyde that his experiments (still in progress) have given confirmatory evidence. On the basis here suggested, the mechanism of such a result is clear: The lens protein does not ordinarily get into the circulation, so the animal neither forms antibodies to it nor binds them when immune serum is injected. These antibodies are, therefore, free in the circulation in the injected rabbits, and some of them combine with the specific genes (in the germ cells) that are responsible for lens antigen—presumably inactivating them. If the results can be confirmed, and if this interpretation is correct, we have here a method for the induction of a specific mutation.

Guyer and Smith interpreted their results as demonstrating the inheritance of an acquired character. On the interpretation here suggested, it becomes a matter of terminology whether they are properly to be so described. It does appear, however, that other cases cited in support of the Lamarckian view should be reëxamined with the gene-antigen-antibody possibility in mind. A casual survey of the literature leads me to suppose that the most likely additional example is that reported years ago by Brown-Sequard. In this case the interpretation might be that guinea pigs are capable of producing antibodies to their own nervous tissue when it is injured.

* This manuscript was written in 1940, and was submitted to Dr. Hyde, whose experiments are referred to as forming much of the basis for the conclusions. Dr. Hyde gave his permission for the reference to his work; but I felt that he would prefer to wait until he could carry out more experiments, and I therefore did not publish. Two factors now influence me to publish the note: Dr. Hyde's death in 1943, without his having published his results; and the results of Dr. Emerson, described in the accompanying paper. The note stands as written in 1940, though a few minor changes might now be desirable. For more recent references, see the paper of Dr. Emerson.

Boyd, W. C., *Tabulae Biologicae*, **17**, 113–240 (1939).

Castle, W. E., and Keeler, C. E., *Proc. Nat. Acad. Sci.*, **19**, 92–98 (1933).

Finlay, G. F., *Brit. Jour. Exper. Biol.*, **1**, 201–213 (1924).

Guyer, M. F., and Smith, E. A., *Jour. Exper. Zool.*, **26**, 65–82 (1920); *Ibid.*, **38**, 449–475 (1924).

Haldane, J. B. S., *Reun. plén. soc. biol.*, pp. 1–14 (1935).

Huxley, J. S., and Carr-Saunders, A. M., *Brit. Jour. Exper. Biol.*, **1**, 215–248 (1924).

Ibsen, H. L., and Bushnell, L. D., *Genetics*, **19**, 293–309 (1934).

Irwin, M. R., *Genetics*, **24**, 709–721 (1939).

Landsteiner, K., *The Specificity of Immunological Reactions*, Charles C Thomas, Springfield, Ill., 178 pp., 1936.

Loeb, L., and Wright, S., *Amer. Jour. Pathol.*, **3**, 251–283 (1927).

Todd, C., *Proc. Roy. Soc. London*, **B117**, 358–366 (1935).

Wiener, A. S., *Blood Groups and Blood Transfusion*, Charles C Thomas, Springfield, Ill., 220 pp., 1935.

THE INDUCTION OF MUTATIONS BY ANTIBODIES*

BY STERLING EMERSON

THE W. G. KERCKHOFF LABORATORIES OF THE BIOLOGICAL SCIENCES, CALIFORNIA INSTITUTE OF TECHNOLOGY, PASADENA, CALIF.

Communicated July 5, 1944

In most biological syntheses each step is controlled by a different gene. This conclusion was suggested by such combined genetic and biochemical studies as those on the anthocyanin pigments[1] from which it appeared that a particular gene regulates the oxidation (or methylation, or the nature of the glucosides) at particular carbon atoms of the anthocyanin molecules. Studies of the eye pigments of Drosophila[2] and the hair pigments of guinea pigs[3] each indicate that the different steps in a particular synthesis are controlled by different genes. Almost overwhelming support for the conclusion is coming from studies of "biochemical" mutations in Neurospora.[4]

Studies of inherited antigenic differences in many species[5] agree in showing that the immunological specificity of an individual antigen is determined by a single gene, as if the development of antigenic specificity represents but a single step in the synthesis of the antigen. The synthesis of genes may represent an analogous instance. "The most fundamental —in fact, the unique and distinctive—characteristic of a gene is . . . the fact that, in its protoplasmic setting, it produces a copy of itself, next to itself, and that when its own pattern becomes changed, the copy it now builds is true to its new self."[6] In the terminology of this report: the specificity of a gene is determined solely by a gene of identical specificity to that synthesized.

The possibility that a gene impresses a copy of its own antigenic specificity on the antigen it regulates can, under sufficiently favorable circumstances, be tested experimentally. Antibodies to such cellular antigens should combine with the genes whose antigenic structure is the same as that of the immunizing antigen. There is no certainty that the gene-antibody reaction can be recognized by the usual serological techniques. There is the chance, however, that if antibodies combine with genes at the time they are "producing copies of themselves," the presence of the antibody molecule would so alter the surface of the gene that the copy produced would have a changed structure. Such alteration in genic structure should, on the basis of this reasoning, result in some change in the physiological or developmental step regulated by the gene in question, and the antibody-induced change should be recognizable as a mutation, as pointed out by Sturtevant in the preceding paper.[7]

While the series of events just postulated presents numerous conditions to be met in an experimental attack, there are certain recorded cases which appear to fit this interpretation. Of these the only one in which antibodies to a relatively pure antigen have induced changes of proved heritable nature is that reported by Guyer and Smith and discussed by Sturtevant.[7] The instances of apparent induction of mutations in bacteria by specific antisera, while suggesting the same interpretation, do not fulfil all conditions since the absence of sexual reproduction in these organisms makes the genetic proof of mutation impossible. Probably the most thoroughly worked out example of changes following treatment with antisera is the smooth to rough transformation in pneumococcus.[8] The smooth, or virulent, form produces a polysaccharide envelope (chemically and immunologically specific to the type) which is absent in the rough form. Changes from smooth to rough rarely occur spontaneously, but are abundantly produced by growth in antiserum to the specific type. While, from the published accounts, it seems probable that the loss of the capacity to synthesize the polysaccharide is directly induced by antibodies to the polysaccharide, it is still impossible to rule out the alternative interpretation that the antibodies act as a powerful selective agent favoring those rough cells which occasionally arise spontaneously.

The ascomycete, *Neurospora crassa*, was selected for the experiments to be reported because of its favorable genetic and cultural characteristics.[9] The general plan of the experiments was to treat the fungus with antisera (developed in rabbits to mycelial extracts or culture filtrates), isolate strains descended from single nuclei (obtained by crossing to the opposite sex and selecting ascospores), and to test the isolates for deficiencies in growth on different carbohydrates.[10] The results of the experiments are summarized in the accompanying table.

To test the genetic homogeneity of the strains used (Beadle's 1A and 25a) conidia of one sex were dusted over cultures of the other sex in which properithecia had already developed. Since the nuclei from one conidium do not ordinarily take part in fertilization in more than one perithecium,[11] it is possible to sample genes from a large number of nuclei of the conidial parent by testing one isolate from each perithecium. The untreated controls listed in the table were handled in this way. Whenever conidia were treated with antisera the isolates were obtained in the same way, but in those instances in which the mycelial stage was treated the method was modified, and there is the possibility that more than one isolate was descended from a single treated nucleus.

The data indicate a much greater variability in the treated material than in untreated controls. No statistical analysis is attempted because the tests for variants were not strictly comparable from one experiment to another. In any one experiment, however, the controls were tested at the same time and in the same way as the treated isolates, and the differences are consistent throughout.

Of the variants listed, eleven have so far been subjected to genetic test, and each of these has differed from the wild-type parents by a single gene. Evidence from genetic tests and from tests of heterocaryons[12] indicates that four variants (Nos. 1077, 1083, 1092 and 1095) from one treatment represent different alleles of a gene by which the parent strains (1A and 25a) differ from another pair of wild-type strains (Beadle's 4A and 12a). At least three other variants represent mutations at three other loci.

MATERIAL TREATED	TREATMENT	IMMUNIZING ANTIGEN	CROSSED TO	TOTAL ISO-LATES	VARIANTS	
					NO.	IDENTIFICATION
Mycelium of 25a	50% antiserum, 27 hrs. at 25°	Mycelial extract	Untreated 1A	10	1	No. 859
	50% antiserum, 27 hrs. at 25°	Culture filtrate	Untreated 1A	21	0	. . .
	50% normal serum, 27 hrs. at 25°	None	Untreated 1A	12	0	(Control)
Conidia of 1A	100% antiserum, 3 hrs. at 25°	Culture filtrate	Untreated 25a	52	4	Nos. 1077, 1083, 1092 1095
	None	None	Untreated 25a	63	0	(Control)
Mycelium of 1A	100% antiserum, 48 hrs. at 25°	Culture filtrate	25a, same treatment as 1A	128	1	No. 1874
	Same, but two 2-day transfers	Culture filtrate	25a, same treatment as 1A	71	2	Nos. 1905, 2606
	Same, but four 2-day transfers	Culture filtrate	25a, same treatment as 1A	210	2	Nos. 2195, 2679
	Same, but five 2-day transfers	Culture filtrate	Untreated 25a	94	1	No. 2012
Conidia of 1A	None	None	Untreated 25a	170	0	(Control)
Conidia of 25a	100% antiserum overnight at 2°	Starch adsorbed enzymes	Untreated 1A	13	2	Nos. 5131, 5168
	100% antiserum overnight at 2°	Starch adsorbed enzymes	Untreated 1A	30	4	Nos. 5175, 5177, 5269, 5294
	100% antiserum overnight at 2°	Amm. sulphate preciptated culture filt.	Untreated 1A	31	3	Nos. 5139 5191, 5201
	100% antiserum overnight at 2°	Amm. sulphate precipitated culture filt.	Untreated 1A	29	5	Nos. 5007, 5008, 5031, 5204, 5211
	None	None	Untreated 1A	31	0	(Control)
			Total treated	695	25	
			Controls	276	0	

With the exception of two morphological mutations, all variants so far obtained differ from the wild type by deficient growth [13] when cultured at standard temperatures and on carbohydrate substrates normally utilized by the wild type. None has yet been obtained which fails to use a particular carbon source under all conditions.

The data just presented indicate that antibodies (or possibly some other serum constituent whose rôle is unsuspected) influence the frequency of recoverable mutations. The indications are that the mutations are induced rather than selected by the treatment, since prolonged treatment of mycelium (during which considerable growth occurs) seems to be less effective than a relatively short treatment of conidia (where few nuclear divisions can have occurred). This difference was especially noticeable when conidia were treated at low temperatures which tend to favor more complete antigen-antibody reactions (e.g., in the precipitian and complement-fixation tests). The apparent deficiency of variants following prolonged growth in antiserum can be interpreted either as a result of selection against the mutated nuclei, or on the assumption that nuclei already affected are more likely to be further affected by antibodies, resulting in inviability.

While the immunizing solutions were known to contain many carbohydrases, the antibodies produced have not been identified in terms of specific antigens.[14] On the present interpretation, the four members of an allelic series may have arisen in response to one sort of antibody, the three independent genes to three other sorts.

The material seems well suited to the problem under discussion, but there are certain refinements to be adopted. Fertile wild-type stocks which lack the growth deficiencies of 1A and 25a, and which give as little variation as possible when intercrossed are being developed. Further, in order to observe a direct relationship between specific genes and antibodies it will be necessary to make use of antigens which can be isolated in relatively pure form. It may be essential to select antigens which are not in themselves indispensable to the organism, otherwise viable mutations are not to be expected.

---

* Work supported in part by a grant from the Rockefeller Foundation. The author wishes to express his indebtedness to Professor G. W. Beadle for the stocks of Neurospora used and for making available unpublished methods developed by him and his collaborators. I am also indebted to my colleagues in this Institution and at Stanford University for many helpful discussions.

[1] References in Scott-Moncrieff, R., *Jour. Genetics*, **32,** 117–170 (1936), and Lawrence, W. J. C., and Price, J. R., *Biol. Rev.*, **15,** 35–58 (1940).

[2] References in Beadle, G. W., and Tatum, E. L., *Amer. Nat.*, **75,** 107–116 (1941), and Ephrussi, B., *Quart. Rev. Biol.*, **17,** 327–338 (1942).

[3] References in Wright, S., *Biol. Symposia*, **4,** 337–355 (1942).

[4] Beadle, G. W., and Tatum, E. L., *Proc. Nat. Acad. Sci.*, **27,** 499–506 (1941); Tatum, E. L., and Beadle, G. W., *Growth, Fourth Growth Symposium*, **6,** 27–35 (1942), etc.

[5] For example, the heritable erythrocyte antigens of man (Landsteiner and others), of fowls (Todd), of rabbits (Castle and Keeler), see references in Wiener, S., *Blood Groups and Blood Tranfusion*, 2nd ed., Charles C Thomas, Springfield, Ill., 1939; references to the Rh-factor in man in Wiener, A. S., and Landsteiner, K., *Proc. Soc. Exp. Biol. Med.*, **53,** 167–170 (1943); erythrocyte antigens of doves, Irwin, M. R., *Genetics*, **24,** 707–721 (1939); of cattle, Ferguson, L. C., Stormont, C., and Irwin, M. R., *Jour. Immunol.*, **44,** 147–164 (1942); the serum protein antigens of doves, Irwin, M. R., and Cumley, R. W., *Genetics*, **27,** 228–237 (1942).

[6] Quoted from Muller, H. J., *Cold Spring Harbor Symposia on Quantitative Biology*, **9**, 290–308 (1941).

[7] Sturtevant, A. H., the preceding paper in these PROCEEDINGS.

[8] Reimann, H. A., *Jour. Exp. Med.*, **41**, 587–600 (1925), has references to earlier literature; Felton, L. D., *Ibid.*, **56**, 13–25 (1932); Paul, J. R., *Jour. Bact.*, **28**, 46–66 (1934). See also discussions in Heidelberger, M., *Amer. Nat.*, **77**, 193–198 (1943), and Avery, O. T., McLeod, C. M., and McCarty, M., *Jour. Exp. Med.*, **79**, 137–158 (1943). A different interpretation of the transformations in pneumococcus is given by Sonneborn, T. M., *Proc. Nat. Acad. Sci.*, **29**, 338–343 (1943).

[9] References in Lindegren, C. C., *Iowa State Col. Jour. Sci.*, **16**, 271–290 (1942), and Beadle and Tatum (1941), loc. cit., footnote 4.

[10] Tests were usually made on sucrose, cellobiose, maltose, starch and α-amylodextrin, on which the wild type grew "normally." The method is essentially similar to that described by Beadle and Tatum (1941), loc. cit., footnote 4.

[11] It has been shown by Backus, M. P., *Bull. Torrey Bot. Club*, **66**, 63–76 (1939), that conidia of *Neurospora sitophila* do not germinate on agar already supporting mycelial growth, but trichogynes from the properithecia attach themselves to such conidia and withdraw the cellular contents.

[12] The method of testing for allelism with heterocaryons is described by Beadle, G. W., and Coonradt, V. L., *Genetics*, **29**, 291–308 (1944).

[13] Growth characteristics were determined by the "tube method" described by Ryan, F. J., Beadle, G. W., and Tatum, E. L., *Amer. Jour. Bot.*, **30**, 784–799 (1943).

[14] Precipitin titres were too low to permit the utilization of an absorption technique.

A RECONSIDERATION OF THE MECHANISM OF POSITION EFFECT

By Boris Ephrussi and Eileen Sutton

Department of Biology, Johns Hopkins University

Communicated June 1, 1944

The phenomenon of position effect has come in recent years to occupy a central position among problems facing the geneticist, in spite of the restriction of well-established cases to one species. This is, no doubt, due partly to the fact that position effect now clearly appears to be much more frequent in Drosophila than originally suspected; but even more to its possible bearing on the problem of the degree of discreteness of the hereditary material. The real significance of position effect in this connection will depend, of course, on the nature of the mechanism which will be found to be responsible for its occurrence; and that is why the clarification of this mechanism appears so urgent.

Speculations concerning the mechanism of position effect have, on the whole, been surprisingly timid. With the exception of the two papers by Muller,[1] and Offermann,[2] especially devoted to the theoretical discussion of this problem and of Schultz's papers[3, 4] on variegation, the literature on position effect, particularly in more recent years, contains very little on the subject, and is seemingly more concerned with providing an exhaustive list of possible explanations than with elaborating one of the interpretations to a point where it becomes capable of experimental verification. As will clearly appear below, the reasons for this reticence are many. Yet it seems useful at this stage to reconsider the various possibilities of explanation of position effect and to compare their relative probabilities after putting them in a somewhat more precise, although speculative, form.

We may first put aside some early interpretations according to which the frequent modification of the action of genes located in the neighborhood of chromosomal breaks is due to mutation of one or more loci or to destruction of genes induced at the time of breakage (Bridges,[5] Muller and Altenburg,[6] Muller[7]). Such interpretations, which really are negations of position effect as such, have since been radically disproved in a series of cases by the demonstrated reversibility of position effect. The possibility remains that some of the changes often designated as position effects (e.g., *Curly*, *Dichaete*, etc.) may really be covered by the above interpretation, and were this demonstrable they would thus be excluded from the category of position effects.

All other and more recently formulated hypotheses as to the mechanism of position effect can be said to belong to one of two essentially different types of interpretation to which we may refer as the kinetic and the structural hypotheses.[8]

Interpretations of the first (kinetic) type can be derived from Sturtevant's remark in his very first paper on position effect that the "influence of the relative position of genes on their effectiveness in development may be interpreted in terms of diffusion and localized regions of activity in the cells."[9] The papers of Offermann[2] and Muller[1] referred to above contain rather detailed discussions of position effect "in terms of diffusion," the argument running somewhat as follows. Genes undoubtedly produce their effects on the phenotype by forming specific substances. These substances must gradually diffuse into the protoplasm, and their distribution must follow a gradient of concentration with a maximum concentration in the immediate neighborhood of their centers of production. Between the moment of their production under the control of the genes concerned and their final action responsible for the phenotypic effect, these substances probably undergo numerous changes, due to interaction with the surrounding protoplasm and with substances involved in other gene controlled reaction chains. If some of the interactions involve the immediate products of the activity of two different genes, their outcome may depend on the distance separating the two genes from which the two interacting products originate.

A slightly different variant of the same scheme is mentioned by Waddington:[10] position effect is pictured here as the result of competition of two genes for the same substrate. It is the distribution of the substrate, not that of the resulting products, which becomes decisive in this case and which follows a gradient in a direction opposite to that described above. Again the occurrence of position effect is assumed to depend on the distance between the "spheres of influence" of the genes concerned. This particular variant of the interpretation is quoted here only because, in this form, it brings us very close to a picture of gene action at present used by Stern, *et al.*, in an interpretation, to be discussed below, of the effects of combinations of *cubitus interruptus* alleles and of position effects at this locus.

Contrary to this first general type of interpretation, the characteristic of which obviously is that it tends to explain position effect on the basis of interactions between gene *products*, or competition between genes for a common substrate, the "structural hypotheses"[1, 11] postulate reversible modifications of the genes themselves or of their structural interrelations. One of the simplest interpretations of this type consists in assuming that chemical bonds unite neighboring genes into integrated, larger units, and that within such a unit, as within any large molecule, changes produced in one of its parts will affect the properties of the whole. On this view position effects are comparable to the steric effects familiar to the organic chemist.

It must be noticed, first of all, that without further understanding of mutations and chromosomal breaks, no sharp line can be drawn between changes of the sort discussed and mutations. As stated above, one can still question whether all changes which have arisen concomitantly with chromosomal breaks and are now classified as position effects really belong to a single class of changes and one may not exclude the possibility suggested by Dobzhansky[11] that "some changes in gene position may induce reversible, and other changes irreversible position effects."

It must be pointed out now that the difficulties of the structural hypothesis in this form have been from the very first so fully realized by its author[1] that he definitely gave preference to what we have called the kinetic hypothesis, as shown by the title of his paper "The position effect as evidence of the localization of the immediate products of gene activity." The following quotation gives Muller's own presentation of the case and, in the last sentence, a possible way out of the difficulty: "A little consideration, however, shows such an interpretation of the 'position effect' of genes to be very implausible, since direct chemical influences of the type in question could hardly extend over distances so vast, from the chemical standpoint, as are those here in question. Certainly the connections between the genes, even if chemical, are of a different nature (allowing of crossing over) from those within the genes, and must allow of the relative chemical independence of the latter: the individual genes, moreover, are very long, from the chemical standpoint, or else lie far apart from one another."... "That, under these circumstances, the direct chemical effect of the substitution of one gene for another should extend, not only to the adjacent gene, so as to affect the latter's function, but even through the latter to one or more genes beyond, is hardly conceivable, unless we postulate some hitherto unfamiliar principle of chemistry, such as might be involved in those forces which are responsible for the synaptic attraction between like genes. The latter are so unusual and act over such relatively great distances that the possibility is not excluded that the same forces which, when acting between genes of identical conformation, cause a harmonious attraction, might, when acting between unlike genes, set up differential stresses on their different parts and so cause distortions in their conformation and consequent changes in their chemical properties."[1] In the latest of Muller's writings[12] on this subject there is rather stronger emphasis on the latter possibility as well as on the striking coincidence of both position effects and somatic pairing in Diptera.

Since the above quoted lines were written the distances over which position effects may extend have been found to be much greater than those implied in Muller's statement. In fact, they are so great in some cases (Demerec,[13]) that, in addition to further emphasizing the impossibility of an explanation of position effect on the basis of steric effects *sensu stricto*, they moreover lead one to question the validity of the kinetic interpretation. The question is: What distance is covered by the assumed gradients of concentration around the genes and how do these distances compare with the distances over which position effects are known to extend? If we reason as above in terms of a consumed substrate, our problem, as suggested by Delbrück in a personal discussion,[14] becomes somewhat similar to that of the adsorption of phage on the surface of bacteria. The latter problem has been discussed by Delbrück[15] from whose paper we quote: "Let us calculate the average density distribution around a resting adsorbing sphere of radius a (bacterium) suspended in a medium that contains initially a uniform density of particles which may stick to the sphere when they encounter it. The average density near the sphere will decrease because the particles are constantly withdrawn by adsorption and soon a stationary density gradient will be set up. The density c as a function of r, the distance from the center of the sphere, is then

$$c(r) = c_\infty(1 - a^1/r)$$

where c_∞ is the average density at a great distance from the sphere, and a^1 is a constant which depends on the probability of adsorption once a particle comes close enough to the surface of the sphere for direct interaction. If every approach to the surface leads to adsorption the density near the surface must drop to zero and we must therefore have $a^1 = a$. If not every approach is successful, the density near the surface will be positive and $a^1 < a$. This stationary density distribution is quite independent of the diffusion constant. It causes, however, a constant flow F towards the surface which is proportional to the density gradient and to the diffusion constant D."

The concentration of the substrate will thus increase centrifugally from the active center according to a hyperbolic curve. In other words, the "sphere of influence" will have no sharp boundaries. However, after a rather rapid initial change of concentration, we will come into a zone where further changes are extremely slow. If we set the limit of the sphere at an arbitrary concentration value of 0.9 (which is a reasonable assumption in view of the fact that there must be in the cell a margin of safety of substrate concentration), the calculation shows that this limit is reached at a distance of about ten times the radius of the active particle (which may be only a part of the gene, in case the gene, like most enzymes, has a prosthetic group) if every encounter of a substrate molecule with the gene is effective. If it is not, this value will be further reduced. Ten times the radius of the active particle is thus, on the above assumptions, our upper limit.

A rather similar result is obtained if the same reasoning is applied to the product P of gene activity (instead of S). For a spherical gene (or active group of the gene), acting at a uniform rate in a homogeneous medium containing no other sources of P, the equilibrium concentration

$$c = c_a \cdot \frac{a}{r}$$

where c_a is the concentration of P at the surface and the other terms the same as above, but in terms of P instead of S. The boundary of the "sphere of influence" (defined as comprising the concentrations of P between 1.0 and 0.1) is at a distance of the same order of magnitude as above, the gradient being in the opposite direction (c decreasing centrifugally).

Let us turn now to position effect. According to Demerec[13] position effect can extend over as far as fifty bands of the salivary gland chromosome. We will assume that a chromonema in the interphase nucleus of a somatic cell, although considerably shorter than the salivary gland chromosome, has the same proportions as the latter. We will visualize the genes as spherical bodies corresponding in position to the dark bands and the intergenic material to the light spaces. The gene then corresponds at the most to $^1/_{100}$ of the total length of the fifty band segment, and position effect can then cover 200 times its radius. By analogy with enzymes it is almost certain that the active part of the gene corresponds to only a small part of the gene molecule and, as it is further possible that one chromosome band contains more than one gene, it follows that the distance covered by position effect in this case is at least 200 times the radius of the gene and probably more.

A different computation in actual Ångström units may be made without assuming a strict similarity between the salivary gland and interphase chromosomes. It is reasonable to assume for the active part of the gene a radius of an order not exceeding 5 Å. The length of the X chromosome at metaphase is about 2 μ. The euchromatic part is approximately 1.4 μ. As the X chromosome has about 1000 bands, 50 bands represents $^1/_{20}$ of this length, or 0.07 μ = 700 Å. The latter value is again an underestimate, because the interphase chromosomes are certainly less coiled, and numerous observations indicate that the genes are active during the interphase. It thus appears that unless we make some further hypothesis which would be unjustified in the present status of our knowledge (assuming, for instance, non-stationary states or interactions of a more complicated nature with substrates or products of other genes, or ascribing particular physiological importance, in the case of interaction of gene products, to the concentration changes outside the boundary of the "sphere of influence"), the "spheres of influence" do not extend far enough to offer an explanation for position effect.

While this reasoning is based on crude assumptions, and cannot therefore disprove the kinetic hypothesis, it is quoted here because it shows that the latter in its simplest form meets with the same difficulty in regard to distance as does the simplest form of the structural hypothesis. It should be pointed out in addition that *position effects are now known to extend over such great distances that it is difficult to see, on the basis of the kinetic hypothesis, why interactions similar to those which occur between two given genes in position effect should not occur between the same genes in their normal location, when the chromosome regions in which they are located happen by chance to lie close together in the nucleus.* Together with the difficulties which the kinetic interpretation encounters when its application is attempted to concrete cases (see below), these arguments reduce considerably its explanatory value. Moreover the italicized statement indicates that the explanation of the mechanism of position effect must be sought in terms of a *factor which will spread along or across the chromosomes rather than through the karyolymph. This characteristic suggests that the factor primarily responsible for position effect corresponds to a change in the physical state of the chromosome itself rather than to a change in the distribution of a substance emanating from it.*[16]

Before we center our attention on any one of the possible aspects of the physical condition of the chromosomes, it may be well to recall the close relationship between the enzymatic and mechanical properties of the muscle protein, myosin, revealed by the recent work of Engelhardt, *et al.*,[17, 18] Needham, *et al.*,[19, 20, 21] Bailey[22] and others. The state of extension of myosin fibers appears to be directly linked to the substrate-enzyme combination, the enzyme (adenosinepyrophosphatase) activity being intimately associated with the fibrous protein itself and the substrate (adenosinetriphosphate) serving as source of the energy required in the process of relaxation. The contraction and extension of the myosin fiber are regarded as results of folding and unfolding of the polypeptid chain, the first of the two processes (folding) perhaps being caused directly by the enzyme-substrate association ("the molecular length of the adenylpyrophosphate being shorter than the length between the groups on the myosin chain to which it becomes attached"[20]).

The association in myosin of fibrous structure with enzyme activity cannot fail to evoke the analogy with the chromosome, with its obviously enzymatic (genic) function and cyclic mechanical changes. Several authors have visualized the chromosome as consisting of a polypeptide backbone, which remains the same throughout its length, and of different linked materials (groups of radicals) responsible for the longitudinal differentiation of the chromosome into somewhat different regions, the genes (compare Demerec[23, 24]).

It is by no means unlikely that in such a structure the state of extension will lead, through changes in spacing, to changes in function. That spatial relationship should play a decisive role in the enzyme substrate combination, the first step of enzyme action, is a safe assumption. That the state of extension (degree of denaturation?) may or may not, according to the case (nature of substance and method of stretching), have profound effects on the specificity and biological activity of a protein seems to be clearly indicated by the heterogeneity of the results obtained by different authors working with various proteins.[25] The important point in the present connection is that in some cases the state of extension *does* affect the biological function, and it may be said that "so long as a biological test for activity is used, a substance which is active after being spread may be active because the change brought about in spreading is reversed during the activity test" (Anson[26]).

Returning now to position effect, it will be noted that the intimate association of chromosomes in somatic pairing, the coincidence of which with the presence of position effect in Diptera has already drawn the attention of Muller[12] and Schultz,[27] may reasonably be assumed to lead to a more precise alignment of the submicroscopic structures than that which prevails in unpaired segments. That synaptic association may actually lead to a stretching out of the whole structure of the chromosome is conceivable and moreover suggested by certain cytological features of the salivary gland chromosomes. (It may be objected indeed that the inference from the pictures of salivary gland chromosomes to the interphase somatic chromosomes is hazardous. At present, however, this seems to be the only possible procedure.) Finally, it is easy to see that the presence of chromosomal rearrangements may lead, through changes in the pairing condition, to modifications of the state of stress in the immediate vicinity of breakage points.

We would like to emphasize that on this scheme in which the change in gene function in position effect is regarded as a result of changed conditions in the chromosome segment concerned, the normal association of the unaffected parts of the chromosome is emphasized as much as, or more than the eventual absence of it in the affected part. Although this may appear as only a minor change of emphasis, it appears important to us (1) because it centers the attention on the established fact of pairing (whatever its mechanism) instead of on the "forces which, in the case of homologous genes, bring about their synaptic association" and which when "acting between unlike genes, should tend to deform them" (Muller[12]); (2) because, if viewed from this angle, the suggested mechanism is not confined in its action to the genes in the chromosome carrying the break. It applies as well to its homologue. The importance of this point will appear when we come to the discussion of a few cases of position effect in the light of the above scheme.

It is clear to us that, in its present form, the scheme offered for the explanation of position effect is too simple to account for all cases which are classified today as position effects. We are aware that it will require at least further refinements, if not remodeling, and we shall refer later in this paper to the difficulties that it encounters at once. First, however, we want to try its application to a few cases which appear particularly relevant.

As pointed out by Stern, to whom we owe the only detailed account of the dosage relationship at a locus which is profoundly affected also by changes in position, these two properties have to be considered together if one wants to test any explanation for the latter. The data of Stern[28] and Stern and Schaeffer[29,30] on the dosage relationships of different wild-type and mutant alleles at the *ci* (*cubitus interruptus*) locus in the normal chromosome 4 have been interpreted by them in terms of overlapping "spheres of influence" of the homologous genes across the chromosomes. The effects obtained could be explained by assuming that different alleles possessed different capacities (1) for combining with a common (limited) substrate, and (2) for converting it into a substance actively contributing to the normal phenotype; so that, for instance, a gene with a high capacity for combining with the substrate and a low efficiency in converting it would interfere with the activity of a more efficient one by depriving it of substrate. Such a hypothesis, as we have seen, can be extended to position effects at the *ci* locus where interference may be assumed between non-homologous genes along the same chromosome, as well as between homologues across the chromosomes. Decrease of dominance of ci^+ in a translocated chromosome [in the combination $R(ci^+)/ci$][31] would thus be due to interference along the chromosome, reducing the efficiency of the ci^+ allele. But in experiments on position effects of the recessive *ci* allele, Stern and Heidenthal[31] have found, first, that the effect of a + allele in the normal chromosome over a translocated *ci* allele ($R(ci)/+$) deviates more from normality than that of a + allele alone (in haplo-4); and second, that the combination $R(ci)/R(ci^+)$ has a more normal phenotype than the combination $R(ci)/+$. While the first of these facts by itself might be explained on the basis of competition across the chromosomes, this interpretation is inconsistent with the second finding, namely, that although the phenotypes of both $R(ci^+)/ci$ and $R(ci)/+$ are less normal than that of $+/ci$, the combination of the two translocations [$R(ci^+)/R(ci)$] partly restores the normal activity instead of decreasing it still more.

Stern and Heidenthal[31] thus reach the conclusion that "many, but not all, facts reported can be interpreted in terms of concepts regarding genic action which have recently been proposed." The same conclusion regarding kinetic interpretations has been derived above from quite a different line of reasoning. This situation obviously requires the introduction of a new parameter, independent of the properties of the genes which were introduced to account for the dosage relations.

The mechanical hypothesis presented here requires no special parameter for the last discussed case as will be seen by its application in the following paragraphs. In Stern's case of heterozygotes for a fourth chromosome rearrangement and a normal fourth [among which $R(ci^+)/ci$ and $R(ci)/+$ are included] the normal pairing of the fourth chromosomes is modified, and the relationships of the chromosomes are such that there may be modification of stress on either the translocated or the normal chromosome, or both, in the region of the break.

As regards $R(ci^+)/R(ci)$ the actual pairing conditions between the two translocations are not known, and this makes it impossible either to predict the phenotype on the basis of the same assumption of modification of stress or to say whether the observed phenotype agrees with this assumption. Indeed, these conditions may be very different from those obtaining in heterozygotes for one translocation, and might be more analogous to those of a translocation homozygote or hemizygote.

The situation shown by data of Dubinin and Sidorov[32, 33] on translocations of ci^+ is much clearer. Any of these translocations which gave rise to position effects when heterozygous ($R(ci^+)/ci$) were found to give a completely normal phenotype when homozygous or hemizygous. *These results suggest that structural heterozygosity might be a necessary condition for the position effect to become effective, and that in the complete absence or complete normality of pairing in the region of the affected locus the position effect is only potential.*

Dubinin and Sidorov also found that flies heterozygous for two different translocations (each separately producing a position effect when heterozygous with a normal *ci* chromosome) were wild type in appearance. It is possible that the chromosome configurations in these heterozygotes correspond in general features to that in the $R(ci^{+})/R(ci)$ combination.

The change of dominance of Punch reported by Oliver[34] is another case in point. Punch is associated with a translocation and is dominant when heterozygous with normal. This effect can be explained as a reduction of the normal activity of the Pu^{+} alleles in both chromosomes, owing to the change in homology along the chromosome, and the resulting change of tension imposed by synaptic attraction towards a different partner on each side of the locus. *Pu* becomes recessive, however, when heterozygous with another translocation (T(2:4)A 34) in which one break is close to the *Pu* locus. The effect of the second translocation is to increase the normal activity of the *Pu* allele in the opposite chromosome, and we should therefore expect to find that the combination of the two translocations facilitates more normal pairing in the *Pu* region.

As Stern and Heidenthal[31] have indicated, an anomaly of the brown-variegated position effects can be eliminated by postulating a factor capable of affecting both chromosomes of a pair. The brown-variegated types are exceptional in that the effect is dominant, and it was previously necessary to assume that the translocated bw^{+} locus is solely responsible for the dominant effect, having acquired a rather exceptional *antimorphic* activity (Muller[7]). On the structural interpretation, however, the bw^{+} allele in the normal chromosome of a heterozygote may also be affected, so that the combined activity of the two alleles may be reduced in some cells below the level required for a wild type phenotype. We would infer, moreover, from the premise that the activity of two bw^{+} alleles is so markedly reduced, that dominant position effects would be obtained even more readily from rearrangements with breaks in the neighborhood of a recessive *bw* allele; for in the $T(bw)/+$ heterozygote we should expect the activity of the + allele to be affected as much as in $T(bw^{+})/+$, while the *bw* allele in the former heterozygote already has a reduced activity as compared with the translocated bw^{+} in the latter. Although *bw* stocks have probably been irradiated rather infrequently, one such position effect is known (bw^{v6}, Glass[35]). The heterozygote $bw^{v6}/+$ shows variegation as a dominant effect as would be expected. Unfortunately no homozygotes for bw^{v6} were obtained before the stock was lost by reversion. Finally, it is of some interest that Schultz[36] found the Plum (dominant brown-variegated) effect to be suppressed in hybrids of *D. simulans* × *D. melanogaster*. Since it had previously been found by Kerkis[37] that synapsis of the salivary gland chromosomes is considerably reduced in these hybrids, the question may be raised whether there is not some causal connection between the two phenomena.

In considering the case of brown variegation it is of interest to notice that, as compared with other position effects, variegated types are characterized by the fact that the change in function of a gene near to a break is expressed to a variable degree in different cells or groups of cells within the same organ. It appears possible to us that this variability will find its explanation on the basis of the same general interpretation here advanced, i.e., as due to variations in stress depending on the particular configurations of the chromosome complement in individual cells. We here find ourselves in close agreement with Muller,[1] when he says: "the degree and the character of gene functioning should also be influenced to some extent by the manner of grouping of the different chromosomes and chromosome regions with regard to each other, i.e., by the (partly accidental) mode of juxtaposition of different portions of the chromonema in the cell. Hence, in different somatic cells of the same individual, there might be some differences in gene functioning, depending upon the way the chromosomes happened to be situated. It is possible that some cases of somatic variegation may be caused in this way. In that case agents, genetic or otherwise, which affected chromosome arrangements, would thereby affect gene expression. A possible case in point might be the influence of the Y chromosome of Drosophila upon somatic variegation."

As the position effects discussed so far are all associated with breaks near the chromocenter, it is of interest to consider one of a somewhat different type. The Bar eye is due to a duplication in tandem of a euchromatic region. At first sight Bar (and the comparable position effects *Confluens Star*, *Hairy-wing*) may appear not to fit the hypothesis, because the phenotype is changed not only in heterozygotes, but also in homo- and hemizygotes. It is known, however, that in the Bar salivary gland chromosome there seems to be a tendency towards synapsis between homologous bands in the two duplicated sections so that even in homo- or hemizygous Bar both the outline of the chromosome and the details of banding are visibly distorted. It is clearly possible that a modification of stress may be occasioned in one part of the chromosome by synapsis between homologous genes along the chromosome as well as by the interchromosomal forces. Thus the difficulty presented by Bar may be only apparent.

A real difficulty must now be considered, however, in connection with the position effects of the w^+ and bw^+ loci. Rearrangements giving position effects of the w^+ locus when heterozygous, still produce the white-mottled phenotype in the homozygous or hemizygous condition (when these are viable), where there is either complete pairing or no pairing at all. The same is true of the brown-variegated homozygotes. Indeed, the latter deviate more from normality than their corresponding heterozygotes[38]

These facts appear to be reconcilable with the above scheme on the assumption that the manifestation of the position effect even in homozygous flies depends on the attachment of the affected loci to heterochromatic regions.

That the pairing conditions of heterochromatic regions differ from those of euchromatin is obvious from their aggregation in the chromocenter. In order to make clearer the possible relation between this so-called "non-homologous" pairing and position effect, a brief discussion of some other properties of heterochromatin is necessary.

Several lines of evidence lead one to suspect that the physical state of the chromocentral regions differs from that of euchromatin. This evidence may be summarized as follows: (1) The frequency of crossing-over per unit cytological length is lower in the neighborhood of the centromeres than elsewhere. This is apparently a function of proximity to the centromere, and not (or not altogether) an intrinsic property of heterochromatin (Mather[39]). It is interesting to note, however, that high temperature increases crossing-over in heterochromatin, irrespective of its position with regard to the centromere, to a level more nearly corresponding to that found in the euchromatin.[39] High temperature also modifies white-mottled and other position effects associated with heterochromatin towards wild-type, and this suggests that the same change in physical state that results in increased crossing-over is also responsible for the increased activity of the loci manifesting position effects. (2) The increase in length of heterochromatic regions in the salivary gland chromosomes as compared with mitotic prophase chromosomes is relatively less than the corresponding increase of euchromatic regions. This might be interpreted as meaning that the proteins of heterochromatic regions are already relatively extended, whereas those of euchromatin are much folded and therefore capable of greater stretching. (3) According to Caspersson[40,41] heterochromatin and euchromatin give different absorption curves in the ultra-violet region of the spectrum. This difference was held to mean that the protein constituents of heterochromatin were histones, while those of euchromatin were partly of the more complex, and more highly folded, globulin type. The validity of this particular interpretation is open to question (Mirsky and Pollister[42]), but if the differences in absorption are real there must be some underlying structural or chemical difference.

Returning now to the "non-homologous" pairing which heterochromatic regions exhibit, it may be pointed out that if the forces responsible for this pairing are operative between any two heterochromatic regions, they must act not only between regions in non-homologous chromosomes, but also between heterochromatic bands in the same chromosome. In this case, pairing conditions must be very similar to those associated with the Bar position effect. It is true that the Bar locus lies within a region of abnormal pairing, while genes concerned in heterochromatic position effects lie outside of such a region, and cannot be directly affected by the conflicting synaptic forces. But if one assumes a difference (for instance, relative inextensibility) in heterochromatic structure on the grounds discussed above, one might expect changes in stress due to non-homologous pairing to be distributed along the chromosome and to affect the folded regions more than the relatively unfolded.

It must be emphasized that this explanation is only tentative, and until experimental proof of such a situation is available, this will remain the most serious difficulty confronting the general interpretation. Even if this difficulty is set aside, we are aware that the interpretation proposed is entirely hypothetical. The justification for presenting it is that it is open to verification in a number of ways, some of which may now be considered.

It has already been mentioned, in referring to some examples of position effect, that without a precise knowledge of their cytology, the results can only be interpreted provisionally. One of the tests to which the hypothesis here presented is subject is a detailed study of the synaptic configurations characteristic of these chromosomal rearrangements. According to the hypothesis, if we compare aberrations which give rise to position effects with those where no phenotypic change is apparent (for example, the two combinations of Punch described by Oliver[34]), we should expect to find that these differ in some respect such as the degree of synapsis or the variability of the synaptic configuration from cell to cell.

It should also be possible to demonstrate position effects across the chromosomes in a number of cases where, as in *R*(*ci*), aberrations are produced with breaks in the neighborhood of recessive alleles. Especially if such an allele were an amorph, a reduction in the activity of the gene in the opposite (normal) chromosome might be expected to result in a deviation from the normal phenotype. The *bw* locus was mentioned as a case in point. Preliminary experiments involving the *w* locus have yielded results which are not incompatible with the interpretation proposed.

A third type of support for the hypothesis might be derived from studying correlatively the effects of different factors on pairing and on the degree of expression of position effects. For instance, non-disjunction may be considered a result of abnormal meiotic pairing; and marked modifications of position effects following non-disjunction have been observed by the junior author. A common causal agent may be responsible for both phenomena, although this has not yet been demonstrated and another interpretation is possible. The above-mentioned absence of pairing in Drosophila hybrids also seems to offer a field for experimentation.[43]

Summary.—The hypotheses on the mechanism of position effect are critically reviewed. An interpretation is offered which regards position effect as a result of chromosome pairing that causes a modification of stress near the affected loci. Several known cases of position effects are dis-

cussed in the light of this interpretation and methods of testing it are suggested.

[1] Muller, H. J., *Fifteenth Int. Physiol. Congr., Leningrad* (1935).

[2] Offermann, C. A., *Bull. Acad. Sci. U. R. S. S., Cl. Sci. Math. Nat.*, **1**, 129–140 (1935).

[3] Schultz, J., *Proc. Seventh Int. Congr. Genet., Edinburgh*, 257–262 (1939).

[4] Schultz, J., *Cold Spring Harbor Symp. Quant. Biol.*, **9**, 55–65 (1941).

[5] In Morgan, T. H., Bridges, C. B., and Sturtevant, A. H., *Bibl. Genet.*, **2**, 1–262 (1925).

[6] Muller, H. J., and Altenburg, E., *Genetics*, **15**, 283–311 (1930).

[7] Muller, H. J., *Proc. Sixth Int. Congr. Genet.*, **1**, 213–255 (1932).

[8] Schultz[3,4] has recently suggested some interesting relationships between nucleic acid synthesis, the balance between euchromatin and heterochromatin in the cell, the position of genes with reference to heterochromatin, and the occurrence of variegation. It seems to us, however, that these relationships have still to be elaborated in much greater detail before they can be coördinated into a general theory of position effect.

[9] Sturtevant, A. H., *Genetics*, **10**, 117–147 (1925).

[10] Waddington, C. H., *An Introduction to Modern Genetics*, Macmillan, New York (1939).

[11] Dobzhansky, Th., *Biol. Rev.*, **11**, 364–384 (1936).

[12] Muller, H. J., *Cold Spring Harbor Symp. Quant. Biol.*, **9**, 151–165 (1941).

[13] Demerec, M., *Genetics*, **25**, 618–627 (1940).

[14] We wish to acknowledge the kind interest and help of Dr. M. Delbrück.

[15] Delbrück, M., *Jour. Gen. Physiol.*, **23**, 631–642 (1940).

[16] According to Dr. Delbrück a linear diffusion along the chromosome might occur, however, if the chromosome had a liquid-liquid phase-boundary from the karyolymph.

[17] Engelhardt, W. A., *Yale J. Biol. Med.*, **15**, 21–38 (1941).

[18] Engelhardt, W. A., Ljubimova, M. N., and Meitina, R. A., *C. R. Acad. Sci. U. S. S. R.*, **30**, 644 (1941).

[19] Needham, D. M., *Biochem. Jour.*, **36**, 113–120 (1942).

[20] Needham, J., Kleinzeller, A., Miall, M., Dainty, M., Needham, D. M., and Lawrence, A. S. C., *Nature*, **150**, 46–49 (1942).

[21] Dainty, M., Kleinzeller, A., Lawrence, A. S. C., Miall, M., Needham, J., and Needham, D. M., *Jour. Gen. Physiol.*, **27**, 355–399 (1944).

[22] Bailey, K., *Biochem. Jour.*, **36**, 121–139 (1942).

[23] Demerec, M., *Amer. Nat.*, **73**, 331–338 (1939).

[24] Demerec, M., *Univ. Penn. Bicent. Conf.* (1940).

[25] References in Lawrence, A. S. C., Miall, M., Needham, J., and Shen, S.-C., *Jour. Gen. Physiol.*, **27**, 233–271 (1944).

[26] Anson, M. L., *Cold Spring Harbor Symp. Quant. Biol.*, **9**, 277 (1941).

[27] Quoted by Muller.[12]

[28] Stern, C., *Genetics*, **28**, 441–475 (1943).

[29] Stern, C., and Schaeffer, E. W., *Proc. Nat. Acad. Sci.*, **29**, 351–361 (1943).

[30] Stern, C., and Schaeffer, E. W., *Proc. Nat. Acad. Sci.*, **29**, 361–367 (1943).

[31] See this issue for details of experiments and symbols. We wish to thank Drs. C. Stern and G. Heidenthal for permitting us to quote their data.

[32] Dubinin, N. P., and Sidorov, B. N., *Amer. Nat.*, **68**, 377–381 (1934).

[33] Dubinin, N. P., and Sidorov, B. N., *Biol. Zhurnal*, **3**, 307–331 (1934).

[34] Oliver, C. P., *Anat. Rec.*, **87**, 33 (1943).

[35] Glass, H. B., *Jour. Genet.*, **28**, 69–112 (1933).

[36] In Morgan, T. H., Bridges, C. B., and Schultz, J., *Yearbook Carnegie Inst.*, No. 36, 298–305 (1937).

[37] Kerkis, J. J., *Amer. Nat.*, **70**, 81–86 (1936).

[38] It is possible, however, that this deviation is somewhat misleading, for the change in phenotype may be broken down into two components—a change from wild type to brown and a change from homogeneous eye color to variegated. The former change is more marked in homozygotes than in heterozygotes, but the latter is diminished.

[39] Mather, K., *Genetics*, **24**, 413–435 (1939).

[40] Caspersson, T., *Naturwiss.*, **31/32**, 514–515 (1940).

[41] Caspersson, T., *Chromosoma*, **1**, 562–604 (1940).

[42] Mirsky, A. E., and Pollister, A. W., *Proc. Nat. Acad. Sci.*, **28**, 344–352 (1942).

[43] This paper was read prior to publication by Drs. M. Delbrück, Th. Dobzhansky, H. B. Glass, B. McClintock, H. J. Muller, C. Stern and several members of this department to whom we wish to express our thanks.

MATERIALS FOR THE STUDY OF THE POSITION EFFECT OF NORMAL AND MUTANT GENES

BY CURT STERN AND GERTRUDE HEIDENTHAL*

UNIVERSITY OF ROCHESTER†

Communicated June 19, 1944

The majority of recent studies on the position effect of genes in *Drosophila* have been concerned with normal alleles of various loci. Different kinds of effects seem to have resulted from change in position. In many cases the phenotypic expression of a normal gene has not been altered by a change of position. Some cases, however, have been observed in which the normal allele assumed a dominant mutant action (e.g., the dominant eye color, Plum). Other observations have indicated that the dominance of a normal allele may be diminished (e.g., cubitus interruptus), or that a normal allele may become more or less inactivated (e.g., white and Notch series). Few generalizations could be made from these results. It was thought, therefore, that some insight might be gained from a study which would attempt to correlate data derived from investigations of change in position of both a normal gene and its mutant alleles. The following represents a first report on such studies.[1]

Material.—The locus chosen was that of cubitus interruptus (*ci*) in the fourth chromosome of *D. melanogaster.* In 1934 Dubinin and Siderov[2] showed that many translocations involving the fourth chromosome cause an unusual change in the effect of the normal allele originally present. While this allele in its new position, both in hemi- or homozygous constitution, still produces a normal phenotype, it causes in heterozygous combination with the recessive mutant allele, *ci*, the appearance of the mutant vein character *ci*, an interruption of the cubital vein.

In order to obtain an unselected sample of translocations involving the fourth chromosome and the mutant allele *ci*, males homozygous for *ci* were x-rayed at 4000 r,[3] and crossed to females containing dominant marker genes in their second and third chromosomes. More than 20,000 F_1 males were individually back-crossed to $ci\ ey^R/ci\ ey^R$ females in order to test for the presence of translocations.[4] Fifty-five translocations involving the fourth chromosome and one or more of the chromosomes *Y*, II or III were found, and stocks of all of them established. Heterozygotes were obtained between all translocations and the allele *ci* in its normal position. About half of the translocations proved to give a more extreme phenotype in combination with *ci* than the typical mutant homozygote *ci/ci*. Five translocations of this group form the basis of the experiments to be reported. These five translocations were selected among those which, during an early stage of these studies, were found to exhibit the position effect described in the following section.

We shall use the symbol *R*(*ci*) for any rearrangement (*R*) which leads to a position effect of the mutant allele *ci*. Correspondingly *R*(+) will be used to denote a rearrangement giving a position effect of a normal allele +. *R*(*ci*) and *R*(+) may be referred to as "position alleles" since they behave as alleles of genes like *ci* and + in the typical position.

Heterozygotes of Position Alleles R(*ci*) *and a Normal Allele.*[5]—It was discovered that at room temperature some of the heterozygous $R(ci)/+^C$ individuals showed slight degrees of the *ci* phenotype, in contrast to $ci/+^C$ heterozygotes which at this temperature range are all normal.[6] In order to obtain reliable data, heterozygotes were raised at a lower temperature, 18–19°, which is known to intensify the expression of the *ci* phenotype. Table 1 shows that all five position alleles *R*(*ci*), in combination with $+^C$, gave a larger number of individuals with the mutant phenotype than did $ci/+^C$ flies. Furthermore, while no individual of the latter consitution deviated more from normality than Class 0, which still shows uninterrupted though thinned veins, many flies of the constitution $R(ci)/+^C$ belonged to the more extreme Classes 1, 2 and 3.

The five translocations differ in the intensity with which they cause the mutant phenotype. For the present, two main groups of *R*(*ci*) alleles will be distinguished: (1) *R*(*ci*) -8 and -48 in which the phenotypic distribution of the heterozygotes $R(ci)/+^C$ deviates rather slightly from that of $ci/+^C$, Class *N* still representing 80 per cent or more of all individuals; and (2) *R*(*ci*) -23, -29 and -32 in which the deviations from normality are larger with only 60 per cent or less of the individuals belonging to class *N*.

Data from a preliminary cytological analysis of these translocations, undertaken by Dr. H. D. Stalker, are included in table 1. Of the two translocations with weak expression one occurred near the chromocenter of chromosome 3 and one near the middle of *3L*, while all three translocations with strong expression were located within close range of each other, about one-third the distance from the distal end of *3R*. All breaks in chromosome 4 presumably lie proximal to 101*F*, i.e., between the centromere and the locus of *ci*, as described for the *R*(+) position alleles by Khvostova and Gavrilova (1935).[7]

Heterozygotes of Position Alleles R(ci) and ci.—As stated above, the five *R*(*ci*) translocations were chosen from those which in the combination *R*(*ci*)/*ci* produce a more extreme *ci* phenotype than the non-translocated *ci*/*ci* homozygote. Specific data permitting a comparison between *R*(*ci*)/*ci* and *ci*/*ci* sibs will be reported elsewhere, but in summary it may be stated that the maximum of the frequency distributions of venation classes in *R*(*ci*)/*ci* are shifted to the next, more extreme class or even farther. Table 2 (right column) shows that the five position alleles differ in the degree to which they cause the expression of the *ci* phenotype in the combination *R*(*ci*)/*ci*. It is possible to distinguish two groups of *R*(*ci*) alleles, one leading to a distribution of phenotypes around a mean located between Classes 1 and 2, and another with the more extreme mean between Classes 2 and 3. Comparing the degree of effect of the *R*(*ci*) alleles in *R*(*ci*)/+ (table 1) with *R*(*ci*)/*ci* (table 3), it is seen that a strong effect in the former constitution is paralleled by a strong effect in the latter and vice versa.

TABLE 1

F_1 Non-ey^D of Wild Type "Canton" ($+^C/+^C$) × $ci\ ey^R/ci\ ey^R$ or $R(ci)/ey^D$* 18°C.

Heterozygous for $+^C$ and:	Break in chromosome 3 near salivary chromosome region:	Phenotypic classes†									
		♀♀					♂♂				
		N	0	1	2	3	*N*	0	1	2	3
ci ey	...	484	1	..	.	..	458	18	..	.	..
R(*ci*) - 8‡	69/70	129	16	5	.	..	157	16	5	.	..
R(*ci*) - 48	Chromocenter	174	10	..	.	..	153	17	..	.	..
R(*ci*) - 23	92 *E*	2	13	21	.	..	40	20	5	.	..
R(*ci*) - 29	94 *A*	21	11	28	4	..	11	26	21	2	..
R(*ci*) - 32	92 *E*	38	33	55	9	..	15	17	58	6	..

* ey^D = eyeless-dominant; chromosome IV.

† *N* = uninterrupted distal section of 4th vein; 0 = uninterrupted but thinned section; 1, 2 = different progressive degrees of interruption; 3 = absence of whole section.

‡ The *R*(*ci*) alleles are numbered in chronological order of their discovery.

Heterozygotes between Two Position Alleles R(+) and R(ci).—A translocation N^{264-20} between the first and fourth chromosome, produced by x-radiation of a fly with normal alleles of *ci*, was obtained through the courtesy of Dr. M. Demerec.[8] Females heterozygous for this translocation show a Notch (N) phenotype. They were crossed to *ci* males. Among the offspring those which were heterozygous for the translocation and *ci* exhibited a typical "Dubinin effect": Their phenotype showed *ci* venation with a maximum at Class 2 (table 3). Thus, the N-translocation may be labeled as possessing a position allele $R(+)$.

Females of the constitution $R(+)/ci$ were mated to males which carried one of the five tested $R(ci)$ alleles and, as a marker for their non-translocated fourth chromosome, the dominant ey^D. The non-ey^D female offspring of these crosses were either $R(ci)/ci$ (table 2, right column) or $R(+)/R(ci)$ (table 2, left column). In all five cases the $R(+)/R(ci)$ flies were more normal than their $R(ci)/ci$ sibs (right column) or than $R(+)/ci$ flies (table 3). Indeed, three different types of $R(+)/R(ci)$ flies were nearly completely normal. Moreover, the curious relation is apparent that those $R(ci)$ alleles which cause an extreme *ci* phenotype in either combination $R(ci)/+$ or $R(ci)/ci$ produce the least extreme expression in $R(ci)/R(+)$.

TABLE 2

F_1 NON-ey^D ♀ ♀ OF $y\ N^{264-20}\ B^iB^i$; $R(+)/ci\ ey^R \times R(ci)/ey^D$ 26°C.

	$R(+)/R(ci)$					$R(ci)/ci\ ey^R$				
	N	0	1	2	3	N	0	1	2	3
$R(ci)-8$	37	12	24	17	1	3	7	27	18	1
$R(ci)-48$	35	12	31	29	4	.	8	45	62	6
$R(ci)-23$	15	1	..	..	.	.	1	1	12	5
$R(ci)-29$	48	1	..	4	.	.	.	4	16	20
$R(ci)-32$	50	..	..	..	.	.	.	3	16	25

TABLE 3

F_1 ♀ ♀ OF $y\ N^{264-20}\ B^iB^i$; $R(+)/y\ Hw\ m^2\ g^4\ dl$-49 * $\times\ ci\ ey^R/ci\ ey^R$ 26°C.

$R(+)/ci\ ey^R$					$Hw/ci\ ey^R$	
N	0	1	2	3	N	0-3
..	4	40	186	27	275	..

* *y*, *Hw*, B^iB^i, m^2 and g^4 are sex-linked factors which segregate together as a unit due to the presence of *dl*-49, an inversion, and of N^{264-20}.

It will be necessary to test different $R(+)$ alleles with the $R(ci)$ alleles recorded in table 2, and with others, in order to find out how general are these unexpected results. We have crossed three of the $R(ci)$ alleles used in combination with the $R(+),N^{264-20}$, with another $R(+),w^{258-18}$. This latter translocation does not show abnormal "Notch" wing characters as does N^{264-20}. Nevertheless, the $R(+),w^{258-18}$, gave equivalent results in combinations with $R(ci)$, indicating that the approach to normal, non-*ci* venation in $R(+)/R(ci)$ is independent of the specific phenotype caused by Notch.[9]

Discussion.—When Dubinin and Siderov[2] found that $R(+)$/haplo-IV and $R(+)/R(+)$ produce normal phenotypes while $R(+)/ci$ flies have interrupted veins, it seemed that such relations between alleles were unique for the position effect. It has since been shown (Stern 1943)[10] that the action of *ci* alleles may be described in terms of combining ability, *c*, with a cellular substrate, *S*, and efficiency, *e*, in transforming *S* into a product, *P*, effective in leading toward a normal phenotype. With the help of these concepts the Dubinin effect can be understood if it is assumed that there is a sufficient amount of substrate and that the $R(+)$ allele has sufficient combining power and efficiency in homo- or hemizygous constitution to cause normality; that in the heterozygote $R(+)/ci$ there is competition for *S* between the $R(+)$ and *ci* alleles, and that the efficiency of *ci* is less than that of $R(+)$, and hence the amount of *P* formed is below normal. Similarly the phenotypic effect of $R(ci)$ in the combination $R(ci)/+$ can be understood as the outcome of competition between $R(ci)$ and + in which the former deprives + of *S* and transforms *S* into *P* at a low efficiency. If the effects of the changed position of the + and *ci* allele are described in terms of dominance and recessiveness, the change from + to $R(+)$ as seen in $R(+)/ci$ versus $+/ci$ would be called a *decrease* of dominance of the + allele while the change from *ci* to $R(ci)$, as seen in $R(ci)/+$ versus $ci/+$ would be an *increase* of dominance of the *ci* allele. The interpretation of position effects in terms of competition resolves the apparent paradox.

TABLE 4

POSSIBLE CHANGES OF GENIC ATTRIBUTES

	COMBINING POWER *c*	EFFICIENCY *e*
1	Unchanged	Lower
2	Unchanged	Higher
3	Lower	Unchanged
4	Higher	Unchanged
5	Lower	Lower
6	Lower	Higher
7	Higher	Lower
8	Higher	Higher

In the preceding paragraph the interaction of the various alleles has been considered separately for certain allelic combinations. This led to the recognition of competition between alleles, and of the relative efficiencies of + as compared with *ci*, of $R(+)$ as compared with *ci*, and of + as compared with $R(ci)$. It is now appropriate to ask in which way the position allele $R(+)$ differs from the original allele +, and $R(ci)$ from *ci*. If the amount of available substrate *S* is the same for the gene in its altered as in its original position, then all changes in effect due to new positions of alleles will be the result of changes in the combining power, *c*, or of the efficiency, *e*, or of both. We shall begin our analysis with this assumption of constancy of *S*. Variability of *S* due to changed position of alleles will be considered later. Since it is possible *a priori* that either *c* or *e* may remain unchanged, or be increased or decreased there are eight combinations possible in which at least one of the two properties is altered (table 4). Considering that there exists competition in the heterozygote $+/ci$ between $+$ and *ci* the decreased effect of $R(+)/ci$ could be brought about by any one of the following changes of the $R(+)$ allele as compared with $+$: (1), (3), (5), (6) and (7). On the other hand, the decreased effect of $R(ci)/+$ could be produced by any one of the following changes of the $R(ci)$ allele as compared with *ci*: (1), (4), (5) and (7). Since it is reasonable to assume that the type of change caused by the altered position is the same, in general, for $R(+)$ and $R(ci)$ three possibilities which account for changes in only one of either $R(+)$ or $R(ci)$, but not in both, can be eliminated, namely (3), (4) and (6). This leaves (1), (5) and (7). These three types of changes all agree in that they involve a lower efficiency of the position allele as compared with the original allele. They differ in the type of *c* attribute which may be unchanged, lower or higher than in the original allele. No information is available at present which permits a decision as to which of these three possibilities is actually realized. Any one of them is not only compatible with the effect of $R(+)/ci$ and $R(ci)/+$ but also with the lowered effect of $R(ci)/ci$ as compared with ci/ci. Competition, with low efficiency of a position allele, accounts not only for the "dominant" mutant effect of $R(ci)$ but may provide a general explanation for the occurrence of dominant effects in loci which otherwise have produced recessive mutants only (e.g., Plum).

It will elucidate the problem further to discuss briefly in what ways a new position may cause a change in the effect of a gene.[11] If genic action consists of a primary interaction with a substrate, *S*, leading to a product, *P*, one may conceive of this interaction as being composed of two separate steps. The first would consist of a reaction leading to a gene-substrate association, the second to restoration of the gene with release of *P*. The term combining power would apply specifically to the first reaction and be equivalent to an association constant. The term efficiency might be equivalent to the velocity constant which characterizes the second reaction, provided the gene products *P* of the different alleles are alike or at least equivalent in their effect. If the gene products are different and have different values in leading toward a given phenotype, then the efficiency, *e*, of an allele is a function of both the velocity constant and of the effectiveness of its specific *P*. The deduction—made under the unproved assumption of constancy of *S*—that the observed position effects at the *ci* locus involve a lowered efficiency, localizes these effects on the second step of the gene-substrate interaction. Whether changes in the first step of this reaction are also involved remains undecided. This analysis of the position effect, however, leaves open the question as to the primary change responsible for the changed reaction. The primary change may be an alteration of actual genic structure, or some kind of alteration due to unusual synaptic forces (Muller, 1935).[12] It may even not involve any alteration of the gene itself but may be due either to a changed local environment under which the reactions must proceed or due to a changed amount of appropriately localized and accessible substrate. If in the new position less substrate were available than originally, a lowered total effect would result in the heterozygote of a position allele and the allele in the normal chromosome. By making, in this case, the unjustified assumption of an unchanged amount of *S* the lowered effect would be mistakenly interpreted as due to changes in the *c* and *e* attributes. An example may clarify this point. Be it assumed that the *R*(+) and *R*(*ci*) positions involve a reduction of available *S*. Since, in *R*(+)/*ci*, *R*(+) and *ci* share in *S* the actual amount of *S* available not only to *R*(+) but also to *ci* might be diminished. Similarly, in *R*(*ci*)/+, the amount of *S* available not only to *R*(*ci*) but also to + might be diminished. If, without knowledge of the decrease in *S*, the amount of the latter, as well as the *c* and *e* properties of the normally located alleles *ci* and + were regarded as unchanged, then the change in total *P* produced by *R*(+)/*ci* and *R*(*ci*)/+ would appear as caused by a lowered *e* of *R*(+) and *R*(*ci*) (with or without change of their *c*). In the absence of information as to the amount of available *S* in heterozygotes for position alleles as compared with typical heterozygotes, we may summarize our analysis by stating that competition between the alleles is involved in the case of position effects of the *ci* locus as it is in typical heterozygotes but that the primary cause or causes of the difference in effect of position heterozygotes as compared with typical heterozygotes remains an open problem. Such a primary factor may either be a diminished amount of available substrate, a decreased efficiency of the position alleles, or possibly a combination of both.

While the proposed interpretation in terms of competition of the position alleles seems to unify most facts concerning gene action of typical and position alleles, it is unsuccessful without additional hypotheses in accounting for the effect of the combination *R*(+)/*R*(*ci*). Since both *R*(+)/*ci* and *R*(*ci*)/+ cause an effect less normal than +/*ci*, it was expected that *R*(+)/*R*(*ci*) would be still less normal. Instead, this constitution causes a shift, in some cases nearly complete, toward normality. Possibly this fact is related to the type of somatic pairing occurring in heterozygotes for two translocations as compared to heterozygotes for only one translocation. No satisfactory interpretation, however, can be proposed at present.[13] It may be added that an apparently similar case has recently been described in an abstract by Oliver (1943),[14] though his interpretation differs from the one given for *ci*. He found that the "dominant" eye color Punch which arose in connection with a translocation becomes "recessive" when combined with another translocation in the homologous chromosome which carries a normal allele.

Incidentally, the data reported furnish a unique example of the reality of "premutation," or "directed mutation" as postulated frequently in the early genetic discussions of evolutionary phenomena. The three "alleles" +, *ci* and *R*(*ci*) according to their effect may be described as dominant normal, recessive mutant and dominant mutant. A dominant mutant allele at the *ci* locus has only twice been found to be due to a single-step mutation from a normal allele. Our data show that such dominant mutant "alleles" can be easily produced by a step involving change of the effect of the recessive mutant *ci* which itself is removed by one mutational step from +.

Summary.—In *D. melanogaster* five translocations between the third and fourth chromosome, carrying the mutant allele *ci*, were studied. In combination with an unbroken fourth chromosome containing a normal allele they produce, in varying degrees, *ci* phenotypes in contrast to the typical heterozygote +/*ci* which appears nearly normal. Combined with *ci* in an unbroken fourth chromosome they lead, in varying degrees, to more extreme *ci* phenotypes than *ci*/*ci*. The three translocations which cause the greatest deviation from normality in combination with + also do so in combination with *ci*. A translocation between chromosomes 1 and 4 derived from a stock containing a normal allele of *ci* produces a striking *ci* phenotype if combined with *ci* in an unbroken chromosome ("Dubinin effect"). If this translocation instead is combined with *ci*, as carried in any one of the five translocations, a shift of phenotypes toward normality occurs. Those translocations which cause the least normal phenotypes in combination with + and *ci* are most effective in causing a shift toward normality when combined with the +-carrying translocation.

Many, but not all, facts reported can be interpreted in terms of concepts regarding genic action which have recently been proposed.[10] Other facts seem to involve additional phenomena. The position effects in heterozygotes appear as competition between the alleles in the normal and changed position.

---

* Now Department of Zoology, Wellesley College.

† Supported in part by a grant from the Rockefeller Foundation.

[1] An abstract had been submitted to the Genetics Society for the 1942 meeting; see Stern, C., *Genetics*, **28**, 92–93 (1943).

[2] Dubinin, N. P., and Siderov, B. N., *Am. Nat.*, **68**, 377–381 (1934).

[3] We wish to acknowledge the kindness of Dr. S. W. Warren, Professor of Radiology, The University of Rochester School of Medicine and Dentistry, who carried out the radiations.

[4] ey^R, a recessive allele of eyeless, served as a marker gene for non-radiated fourth chromosomes. It is without influence on the expression of *ci*. Except in the tables, reference to ey^R will be omitted.

[5] The normal allele was the $+^C$ allele of the *ci* locus, see Stern, C., and Schaeffer, E. W., *Proc. Nat. Acad. Sci.*, **29**, 361–367 (1943).

[6] After completion of the manuscript of the present communication a paper by Siderov, B. H., *Compt. rend. acad. sci. U.R.S.S.*, **31**, 390–392 (1941), has come to attention in which similar effects of *R*(*ci*) alleles in combination with + are described.

[7] Khvostova, W. W., and Gavrilova, A. A., *Biol. Zhurnal.*, **4**, 905–916 (1935).

[8] "Salivary chromosomes (Sutton) show break in *X* at 3 C, with deficiency of 3C 5–7; and break in 4 at 101*F*," Bridges and Brehme, "The Mutants of *Drosophila melanogaster*," *Carnegie Inst. Wash. Publ.*, **552** (1944).

[9] In the paper of Siderov referred to in footnote 8 a further remarkable fact is reported. Two translocations between the third chromosome and a fourth chromosome which carries a + allele but which do not give a Dubinin effect in combination with *ci*, shift the phenotype toward normality if combined with *R*(*ci*).

[10] Stern, C., *Genetics*, **28**, 441–475 (1943).

[11] Grateful acknowledgment is made to Dr. D. R. Goddard for clarification of some of the concepts involved in this discussion.

[12] Muller, H. J., *Summ. Commun. XV int. physiol. Congr.* (Leningr.-Mosc.), pp. 286–289 (1935); and *Proc. 15th int. physiol. Congr.* (Leningr.-Mosc.), pp. 587–589 (1938).

[13] If it were true that the new positions imply a reduction of *S* at the *ci* locus and if heterozygotes between normally located and position alleles frequently undergo somatic pairing while compounds between *R*(+) and *R*(*ci*) rarely pair, then competition might take place between the paired alleles of heterozygotes, while no competition might occur between the separated alleles *R*(+) and *R*(*ci*). The lack of competition might then permit a shift toward normality. The assumptions made are liable to experimental test.

[14] Oliver, C. P., *Anat. Rec.*, **87**, 33 (1943).

COMPARISON OF A DEGENERATE FORM OF EINSTEIN'S WITH BIRKHOFF'S THEORY OF GRAVITATION

BY HERMANN WEYL

INSTITUTE FOR ADVANCED STUDY, PRINCETON, NEW JERSEY

Communicated July 20, 1944

Whereas electric charge is not universally proportional to the inertial mass of bodies their gravity is. This fundamental fact, supported both by daily experience and the most refined experiments, led Einstein to the conception that inertia and gravitation are one (principle of equivalence) and thus to his theory of general relativity. The main reason for my and many others' belief in that theory is the radical explanation it affords for the fact just mentioned. Any theory which breaks up the unity of inertia and gravitation, as Birkhoff's recent theory of gravitation in a flat world[1] does, throws us back into the position before Einstein where we had to accept the identity of mass and weight without understanding it. Is there any reason for such a withdrawal?

One way of breaking the bond between inertia and gravitation is suggested by the study of weak gravitational fields in Einstein's theory: One lets the g_{ik} differ but infinitesimally from constant values δ_{ik}, $g_{ik} = \delta_{ik} + \epsilon h_{ik}$ (δ_{ik} = inertia, h_{ik} = gravitation), throws away all higher powers of the constant factor ϵ except the first, writes h_{ik} for ϵh_{ik} and adopts the equations thus obtained as strictly valid. In a review[2] of Birkhoff's paper I had tried to illustrate his scheme by analogy with this "degenerate Einstein theory." In a recent note[3] A. Barajas claims to disprove the validity of this comparison. But he misinterprets my intention, and his calculations are not entirely correct. I hope to clarify the situation by the following fuller statement of the case, in which I compare the complete and the "degenerate" Einstein theories (E, D) with the theory of Birkhoff (B).

In terms of a suitable parameter σ the equations of the world trajectory of a particle have the common form

$$\frac{d^2x_i}{d\sigma^2} = -\gamma\cdot\Gamma^i_{pq}u^p u^q \qquad (u^i = dx_i/d\sigma), \tag{1}$$

which is invariant with respect to the substitution $\sigma \rightarrow k\sigma$. The components Γ^i_{pq} of the gravitational field are in the respective theories defined by

$$\Gamma_{i,pq} = g_{ij}\Gamma^j_{pq} = \frac{1}{2}\left(\frac{\partial g_{ip}}{\partial x_q} + \frac{\partial g_{iq}}{\partial x_p} - \frac{\partial g_{pq}}{\partial x_i}\right), \tag{2E}$$

$$\Gamma_{i,pq} = \delta_{ij}\Gamma^j_{pq} = \frac{1}{2}\left(\frac{\partial h_{ip}}{\partial x_q} + \frac{\partial h_{iq}}{\partial x_p}\right) - \frac{\partial h_{pq}}{\partial x_i}, \tag{2B}$$

$$\Gamma_{i,pq} = \delta_{ij}\Gamma^j_{pq} = \frac{1}{2}\left(\frac{\partial h_{ip}}{\partial x_q} + \frac{\partial h_{iq}}{\partial x_p} - \frac{\partial h_{pq}}{\partial x_i}\right). \tag{2D}$$

The constant factor γ, specific for the individual particle, is the ratio gravity/mass. Once the cut performed by which D arises from E, a combination $\delta_{ik} + \mu h_{ik}$ with this or that constant μ has as much and as little significance in D as in B. In B and D there is no more reason for γ to be independent of the particle than one can expect the ratio η = charge/mass to be the same for all particles moving in an electromagnetic field according to the equations

$$\frac{d^2x_i}{d\sigma^2} = \eta\cdot F^i_q u^q \qquad \left(F_{ik} = \frac{\partial\varphi_k}{\partial x_i} - \frac{\partial\varphi_i}{\partial x_k}\right).$$

We use the familiar notations

$$ds^2 = g_{ik}dx_i dx_k \text{ in E,}$$

$$ds^2 = \delta_{ik}dx_i dx_k = dx_0^2 - (dx_1^2 + dx_2^2 + dx_3^2) \text{ in B and D.} \qquad (3)$$

From (2E) and (2B) there follows the relation

$$\left(\frac{ds}{d\sigma}\right)^2 = \text{const.} \qquad (4\text{E, B})$$

Birkhoff's choice of the Γ^i_{pq} was, in fact, suggested by the postulated relation (4B). No such relation (4) holds in D. [One could remedy this by a suitable transformation of the parameter σ before letting $\epsilon \rightarrow 0$; the recipe for the construction of D is not entirely unambiguous! But following Mr. Barajas I shall stick to our present interpretation of D.]

What I had in mind when I said that B and D are "essentially of the same nature" (not "much the same," as Mr. Barajas quotes me) were the following points: (1) Both operate in a flat world and are linear with respect to h_{ik}. (2) Both have a *tensor* h_{ik} as gravitational potential, and the ponderomotoric force exerted by the gravitational field depends *quadratically* on u^i. One should not forget that before Einstein a gravitational potential with 10 components was an unheard-of thing; this complication was forced upon him by the principles of general relativity rather than by any observed facts about the dependence of the force of gravitation on velocity. Notice also the formal similarity of the expressions (2B) and (2D). (3) In both theories inertia and gravitation are separate entities; hence no account can be given for the identity of weight and mass, while the equation $\gamma = 1$ is a necessity in E—but pointing to these analogies I did not for a moment suggest that B and D are identical in their numerical consequences!

In a theory of discrete particles the static centrally symmetric solutions of the homogeneous field equations are decisive. With such calibration of the distance $r = (x_1^2 + x_2^2 + x_3^2)^{1/2}$ from the center as renders Einstein's metric 3-space a conformal map of Euclidean space this solution in E is given by

$$ds^2 = \left(\frac{1 - m/2r}{1 + m/2r}\right)^2 \cdot dx_0^2 - (1 + m/2r)^4(dx_1^2 + dx_2^2 + dx_3^2). \quad (5\text{E})$$

It contains only *one* constant m. In the limit for small m/r one obtains

$$h_{00} = -2m/r, \qquad h_{0\alpha} = 0, \qquad h_{\alpha\beta} = (2m/r)\delta_{\alpha\beta} \quad [\alpha, \beta = 1, 2, 3]. \, (5')$$

The homogeneous field equations assumed by Birkhoff are

$$\Box h_{ik} = 0 \qquad \text{where } \Box\varphi = \delta^{pq}\frac{\partial^2\varphi}{\partial x_p \partial x_q}. \qquad (6)$$

Their most general static centrally symmetric solution involves 3 arbitrary constants a, b, l:

$$\begin{aligned} h_{00} &= -2a/r, \qquad h_{0\alpha} = 0; \\ h_{\alpha\beta} &= (2b/r)\delta_{\alpha\beta} + l(r^{-3}\delta_{\alpha\beta} + 3r^{-5}x_\alpha x_\beta). \end{aligned} \tag{5B}$$

From the present standpoint this is a serious disadvantage of B. (5′) is contained in (5B) as the special case $a = b$, $l = 0$. To have a common basis for the computation of the consequences of (2B) and (2D) I assume the h_{ik} to be given by (5B) with $l = 0$, and set $\gamma = 1$ for all particles.

Then one finds for the period of revolution T of a planet around the sun and for the advance ω of its perihelion in the three competing theories the approximate formulae

$$\begin{aligned} &T^2 \sim 4\pi^2 r_0^3/M, \qquad \omega \sim 2\pi j/r_0(1 - e^2); \\ &M_E = m, \qquad M_B = 2a, \qquad M_D = a; \\ &j_E = 3m, \qquad j_B = 2a + 4b, \qquad j_D = 2a + 3b. \end{aligned}$$

$r_0(1 - e)$ and $r_0(1 + e)$ are the limits between which the distance of the planet from the center (sun) varies. In carrying out the calculation for D, Mr. Barajas changes at a certain step b into $-b$ and erroneously makes use of the relation (4B) which does not hold in D. In our notations, and with the dot used as an abbreviation of $d/d\sigma$, the correct equations in D read

$$\ddot{x}_0 = -\frac{2a}{r^2}\dot{x}_0\dot{r}$$

$$\ddot{x}_\alpha = -\frac{x_\alpha}{r^3}\{a\dot{x}_0^2 + b(\dot{x}_1^2 + \dot{x}_2^2 + \dot{x}_3^2)\} + \frac{2b}{r^2}\dot{x}_\alpha\dot{r},$$

from which, after proper normalization of the constant factor k in σ and of the invariant plane, follow the integrals:

$$\dot{x}_0 = e^{2a/r},$$

$$x_3 = 0, \qquad x_1\dot{x}_2 - x_2\dot{x}_1 = C \cdot e^{-2b/r},$$

$$\dot{x}_1^2 + \dot{x}_2^2 = \frac{a}{2a + b}(e^{4a/r} - e^{-2b/r}) + C_0 \cdot e^{-2b/r}.$$

For the deflection 2δ of a light ray whose asymptote has the distance r_0 from the center one finds

$$\delta \sim i/r_0; \qquad i_E = 2m, \qquad i_B = 2(a + b), \qquad i_D = a + b,$$

if one assumes that the light ray follows a trajectory which satisfies the condition $(ds/d\sigma)^2 = 0$ in E and B and the same condition *at infinity* in D. It is true that the special choice $a = b = m/2$ reduces the *three* values of M, j, i in B to those in E, because the linear relation $M + j = 2i$ satisfied in E holds also in B, whereas $M + j = 3i$ in D.

However, the hypothesis on which our computation of δ is based is of a very dubious nature in B and D. In truth the propagation of light is determined by Maxwell's electromagnetic field equations, from which the concept of light ray emerges by letting the frequency of the light waves tend to infinity. In Einstein's general relativity theory the g_{ik} affect Maxwell's equations in such a way that the light rays thus defined coincide automatically with the nul geodesics; and the deflection $\delta = 2m/r_0$ results. But in B with its flat world there is no necessity for modifying Maxwell's equations, and hence the light rays are straight lines ($i_B = 0$) and do not coincide with the nul trajectories—unless the discrepancy is removed by some artificial modification of the electromagnetic equations which has to be constructed *ad hoc.*

Finally we discuss the relative shift $\Delta\lambda/\lambda$ of wave lengths λ which is observed by a spectroscopist who compares the light emitted from two atoms of the same nature. In Einstein's theory there is no doubt about the principles on which the computation of $\Delta\lambda/\lambda$ rests: the constitution of the atom determines the frequencies of its oscillations measured in terms of its *proper time s*, propagation takes place in spherical waves according to the electromagnetic equations. For the static field (5E), two resting atoms at fixed distances r and a resting observer one thus obtains $\Delta\lambda/\lambda \sim m \cdot \Delta(1/r)$. Desperate steps must be taken to escape the conclusion $\Delta\lambda = 0$ in B. If I do not misunderstand him, Mr. Barajas interprets $dx_0/d\sigma$ as the energy of a photon traveling along its nul trajectory from one to the other atom and, by grafting Planck's quantum law on his classical scheme, assumes that the frequency of the photon varies in proportion to this its energy! In this way (to which the quantum theory of light lends little support) he manages to obtain a red shift $\Delta\lambda/\lambda \sim 2a \cdot \Delta(1/r)$.

It is well known that in E the formula (4E) is nothing but a special case of the law that the scalar product of two vectors v^i, w^i at a common world point P does not change if both undergo parallel displacement to an infinitely near point:

$$dv^i = -\Gamma^i_{pq} v^p dx_q, \qquad dw^i = -\Gamma^i_{pq} w^p dx_q$$

imply $d(v^i w_i) = 0$. Perpendicular vectors remain perpendicular. Among other applications of the latter fact I mention Fokker's computation of the geodesic precession of the Earth's axis,[4] for which, assuming the orbit to be circular, he finds the value $3\pi m/r_0$. How would this phenomenon be treated in B? Birkhoff's rule (2B) naturally fails to produce a law of constancy of the same generality; and yet the special case where v and w are both parallel to the displacement is made the decisive postulate.

When one passes from discrete particles to continuous distribution of charge, energy and momentum and thus from the homogeneous to non-homogeneous field equations, the divergences between the theories E and B become more pronounced than their analogies, for the simple fact that Birkhoff's theory is based on the duality of field and matter in a form which physics has abandoned during the last 40 years. In his scheme mechanical equations appear independently besides field equations, and matter is represented by a charge vector s^i and an energy-momentum tensor T^{ik} of the type

$$s^i = \rho u^i, \qquad T^{ik} = \mu u^i u^k \tag{7}$$

which involve a velocity field u^i (hydrodynamical models of atoms!). In a steady development physics has come to the conclusion that the mechanical equations, the differential conservation laws for charge, energy and momentum, follow from the field equations, and in formulating its exact field laws it has learned to dispense with quantities like (7) which are tied to the idea of a "fluid," a continuously distributed moving substance. It is true that quantum physics by its statistical interpretation of the field equations gives rise to a certain complementary duality of "wave" and "particle"; but this is an altogether different affair.

A consistent linear field theory of gravitation is well suited to illuminate the problem of energy and momentum; it might also help, though this is much more problematic, in pointing the way to the correct unification of gravitation and electromagnetism. I propose, therefore, to continue this discussion at another place.

---

[1] Birkhoff, G. D., these PROCEEDINGS, **29,** 231 (1943).

[2] *Mathematical Reviews,* **4,** 285 (1943).

[3] These PROCEEDINGS, **30,** 54 (1944).

[4] *Versl. K. Akad. Wetensch. Amsterdam,* **29,** p. 611; Schouten, J. A., *ibid.,* p. 1150; Weyl, H., *Raum Zeit Materie,* 5th ed., 1923, p. 320.

SCALE CURVES IN GENERAL CARTOGRAPHY*

BY EDWARD KASNER AND JOHN DE CICCO

DEPARTMENTS OF MATHEMATICS, COLUMBIA UNIVERSITY AND ILLINOIS INSTITUTE OF TECHNOLOGY

Communicated June 29, 1944

1. The square of the linear element dS of a surface Σ is

$$dS^2 = E(x, y)dx^2 + 2F(x, y)dxdy + G(x, y)dy^2, \qquad (1)$$

where $H^2 = EG - F^2 > 0$. Of course (x, y) are general curvilinear coördinates of any point on Σ. Also we may consider (x, y) to represent the cartesian coördinates of any point in a plane π so that the square of the linear element ds on π is $ds^2 = dx^2 + dy^2$. Thus there results a mapping of a surface Σ upon a plane π such that points of Σ and π will correspond if and only if they have the same coördinates. We shall term any particular mapping of a surface Σ upon a plane π a *cartogram*.

We consider the ratio σ of the differentials of the arc lengths on π and Σ, respectively. This scale function σ is defined by the formula

$$\sigma^2 = \left(\frac{ds}{dS}\right)^2 = \frac{1 + y'^2}{E + 2Fy' + Gy'^2} \qquad (2)$$

where the accent denotes differentiation with respect to x.

A scale curve is the locus along which the scale does not vary. There are, in general, ∞^2 scale curves. We shall be chiefly concerned with systems of scale curves, the discussion of which begins in Section 3.

The scale function σ depends upon the lineal-element (x, y, y') only. It is independent of y' if and only if the mapping of Σ upon π is conformal. It will be a mere constant if Σ is a developable surface and the mapping is an unrolling of Σ upon π followed by a similitude on π. Since we have already studied the simple-infinity of scale curves in conformal maps,[1] we shall assume in the present paper that the mapping is not conformal. Then the conditions $F = 0$ and $E = G$ cannot hold simultaneously.

Our new general theory will be applied to classic map projections or cartograms (area-preserving, azimuthal, etc.) in later papers.

2. If on the plane π, the scale function σ is laid off along the corresponding direction y' at a fixed point (x, y), the end-point will describe a locus, which is an ellipse. Let (X, Y) represent the cartesian coördinates of a point on the ellipse relative to the fixed point (x, y), which is, of course, the center of the ellipse. The equation of this terminal ellipse is

$$EX^2 + 2FXY + GY^2 = 1. \qquad (3)$$

The directions of the major and minor axes of this terminal ellipse are given by

$$Fy'^2 + (E - G)y' - F = 0. \tag{4}$$

This defines, according to a theorem of Tissot, the unique orthogonal net of curves on π which by the mapping corresponds to an orthogonal net on Σ. A curve is a characteristic curve of the mapping of Σ upon π if and only if it is a curve of this net.

The semi-major and semi-minor diameters of the terminal ellipse satisfy the equation

$$H^2\sigma^4 - (E + G)\sigma^2 + 1 = 0. \tag{5}$$

For a fixed σ, the preceding equation defines a curve such that at all points of this curve the corresponding terminal ellipses are congruent. This set of $2\ \infty^1$ curves is termed the *singular family* of curves.

(It is noted that this singular family can consist of ∞^1 curves if and only if, either, the mapping is conformal; or else the functions $E + G$ and $H^2 = EG - F^2$ are functionally related. For example, under any area preserving map the singular family is $E + G = \text{const.}$)

3. Upon setting the scale σ of the equation (2) equal to a constant, there result curves along which the scale is constant. *If the mapping is not conformal, there are ∞^2 scale curves.* On the other hand if the mapping is conformal, there are ∞^1 scale curves. Obviously, all curves are scale curves if Σ is a developable surface and the mapping is an unrolling of Σ upon π followed by a similitude on π.

Henceforth only the general case where the mapping is not conformal shall be considered. It is observed that for a fixed σ, there pass two scale curves through any point of the plane. However, at any point of the curve (5), the two scale curves will be tangent. This supplies us with the reason for calling the family (5) singular.

Upon eliminating the constant σ from the equation (2) by differentiation, we discover that *the differential equation of our ∞^2 scale curves is*

$$y'' = \frac{(1 + y'^2)[E_x + y'(E_y + 2F_x) + y'^2(2F_y + G_x) + y'^3G_y]}{2[-F + y'(E - G) + y'^2F]} \tag{6}$$

It is seen that not every system of ∞^2 curves can represent the scales of a general mapping of a surface Σ upon a plane π, because (6) is of special algebraic form in the first derivative.

We observe from (6) that there are *three inflectional directions and two cuspidal directions for the scale curves. The cuspidal directions are always orthogonal.*

4. Let (X, Y) represent the cartesian coördinates relative to the point (x, y) of the center of curvature constructed at the point (x, y) of a scale curve. For a fixed point (x, y), the corresponding centers of curvature (the central locus) of the ∞^1 scale curves through the fixed point describe the cubic curve

$$G_yX^3 - (2F_y + G_x)X^2Y + (E_y + 2F_x)XY^2 - E_xY^3 + 2[-FX^2 + (E - G)XY + FY^2] = 0. \quad (7)$$

The locus of the centers of curvature (the central locus) is a general cubic curve which has a node at the fixed point (x, y), the directions of the tangent lines at the node coinciding with the characteristic directions.

This property is possessed not only by any family of ∞^2 scale curves but also by a more general class of families of curves. If a system of curves possesses the preceding property, then it must be defined by a differential equation of the form

$$y'' = \frac{(1 + y'^2)(\alpha + \beta y' + \gamma y'^2 + \delta y'^3)}{2(-\eta + \epsilon y' + \eta y'^2)}. \quad (8)$$

This family will represent scale curves if and only if the six functions $(\alpha, \beta, \gamma, \delta, \epsilon, \eta)$ of (x, y) satisfy the two partial differential equations of second order

$$\begin{aligned} &M_y = N_x, \\ &\alpha_y + N\alpha = \delta_x + \epsilon_{xy} + N\epsilon_y + \epsilon N_y + M(\delta + \epsilon_y + \epsilon N), \end{aligned} \quad (9)$$

where M and N are defined by

$$\begin{aligned} (\epsilon^2 + 4\eta^2)M &= \epsilon(\alpha - \gamma) + 2\eta(\beta - \delta) - 4\eta\eta_x + 2\epsilon\eta_y - \epsilon\epsilon_x - 2\eta\epsilon_y, \\ (\epsilon^2 + 4\eta^2)N &= -2\eta(\alpha - \gamma) + \epsilon(\beta - \delta) - 2\epsilon\eta_x - 4\eta\eta_y + 2\eta\epsilon_x - \epsilon\epsilon_y. \end{aligned} \quad (10)$$

There are essentially ∞^2 surfaces which possess the same family of ∞^2 scale curves. If the surface Σ with the linear element (1) is mapped upon the plane π such that the differential equation (6) represents the scale curves, then the linear element dS' of any surface Σ' with the same scale curves must be given by

$$dS'^2 = a dS^2 + b ds^2 \quad (11)$$

where $a \neq 0$ and b are constants.

It must be emphasized that the results of this section are valid only for the general case where the cubic curve (7) is not degenerate.

The class of scale curves is converted into itself only by the conformal group. This is true also for the class of curves defined by differential equations of the type (8).

5. If the parametric curves of Σ are not orthogonal, that is, $F \neq 0$, the cubic curve (7) may degenerate into three straight lines, two of which are the tangent lines of the characteristic curves at the fixed point, and the other is the general line

$$G_yX + E_xY - 2F = 0. \quad (12)$$

In this event, the ∞^2 scale curves form a velocity system.[3] The necessary and sufficient conditions that the ∞^2 scale curves (6) form a velocity family are

$$\begin{aligned} G_y(E - G) &= F(E_x + G_x + 2F_y), \\ -E_x(E - G) &= F(E_y + G_y + 2F_x). \end{aligned} \tag{13}$$

The differential equation of the ∞^2 scale curves is then

$$y'' = (1/F)(1 + y'^2)(-E_x + y'G_y). \tag{14}$$

If the scale curves form a natural family, then the characteristic curves are an isothermal net. The curves $(E - G)^2 + 4F^2 =$ const. form an isothermal family if the scale family is isogonal. The scale family is a conformal rectilinear wex if and only if the function $\log [E - G - 2iF]$ is an analytic function of $z = x + iy$ only. Therefore if the scale family is a conformal rectilinear wex, the curves $E - G =$ const. and $F =$ const. form an isothermal net.

A system of ∞^2 scale curves can be of the cubic (Lie-Liouville) type[3] *if and only if it is a velocity system.*

6. If the parametric curves are orthogonal, so that $F = 0$, the cubic curve (7) can be degenerate if and only if $G_y = 0$ or $E_x = 0$. In this event, the cubic curve consists of the tangent line to one of the characteristic curves and a conic section tangent to the other characteristic curve at the given point. This may also happen in the case where the parametric curves are not orthogonal.

If $G_y = E_x = 0$, together with the condition $F = 0$, the ∞^2 scale curves form the velocity system

$$y'' = \frac{1}{E - G}(1 + y'^2)(E_y + G_x y'). \tag{15}$$

In this case, the cubical central locus consists of the tangent lines of the characteristic curves at the fixed point, and the line

$$-G_x X + E_y Y + 2(E - G) = 0. \tag{16}$$

7. Finally we determine the class of surfaces Σ for which there exists a map of Σ upon the plane π such that the scale curves coincide with the totality of α^2 straight lines. This class of surfaces Σ is given by

$$\begin{aligned} dS^2 = a_0(ydx - xdy)^2 + 2(a_1dx + b_1dy)(ydx - xdy) + a_2dx^2 \\ + 2c_2dxdy + b_2dy^2. \end{aligned} \tag{17}$$

For this class of surfaces, we find

$$\begin{aligned} H^2 = (a_0a_2 - a_1^2)x^2 + 2(a_0c_2 - a_1b_1)xy + (a_0b_2 - b_1^2)y^2 + 2(a_1c_2 - a_2b_1)x \\ + 2(a_1b_2 - b_1c_2)y + (a_2b_2 - c_2^2). \end{aligned} \tag{18}$$

The gaussian curvature G is given by

$$H^4G = -2a_0H^2 + a_0(c_2^2 - a_2b_2) + a_2b_1^2 - 2a_1b_1c_2 + a_1^2b_2. \qquad (19)$$

Thus the guassian curvature G is found to be a rational function of (x, y).

In the general case (17), it is found that the ∞^1 straight scales corresponding to a given scale are tangent to a conic section. By varying σ we find that the conic sections form a confocal family. It is a central conical or parabolic family according as $a_0 \neq 0$ or $a_0 = 0$.

If the gaussian curvature G is identically constant, then G must be identically zero. In this case $a_0 = a_1 = b_1 = 0$. This means that if the surface Σ is of constant curvature such that the scale curves are straight lines, then Σ must be developable, and the mapping of Σ upon π is a development of Σ upon π followed by an affine transformation. To a given scale σ, corresponds a parallel pencil of straight lines.

Thus there exists no non-conformal map of a sphere upon a plane such that the scale curves coincide with the totality of ∞^2 straight lines. From this, we deduce that the only map (conformal or not) of a sphere upon plane such that the scale curves are straight lines is the Mercator projection (in which case we have a single-infinitude of parallel straight lines as scale curves).

---

[*] Presented to the American Mathematical Society, February, 1944.

[1] See the paper, "Scale Curves in Conformal Maps," *Proc. Nat. Acad. Sci.*, (1944). Also see "Scale Curves in Cartography," *Science*, **98**, 324–325 (1943), and *Science News Letter*, March 25 (1944).

[2] A general velocity system is defined by the differential equation $y'' = (1 + y'^2)(\psi - y'\phi)$ where ϕ and ψ are functions of (x, y). See Kasner, "Differential Geometric Aspects of Dynamics," *Amer. Math. Soc. Colloquium Publications*, **3** (1912, 1934); *Trans. Amer. Math. Soc.*, **9**, **10** (1908–1909). See also Kasner and De Cicco, "The Geometry of Velocity Systems," *Bull. Amer. Math. Soc.*, **49**, 236–245 (1943), and Kasner, "Geometric Properties of Isothermal Families," *Instituto de Matematica*, Rosario, Argentina, 5 (1943), volume in honor of Rey Pastor.

[3] A differential equation of the cubic type is of the form $y'' = \alpha + \beta y' + \gamma y'^2 + \delta y'^3$, where $(\alpha, \beta, \gamma, \delta)$ are functions of (x, y) only. This has been studied extensively by Lie, R. Liouville, Tresse, Kasner and Wilczynski. With respect to the group of arbitrary point transformations, this type is called a differential equation of the first rank. See Kasner, "The Geometry of Differential Elements of the Second Order with Respect to the Group of All Point Transformations," *Amer. Jour. Math.*, **28**, 203–213 (1907).

PROCEEDINGS
OF THE
NATIONAL ACADEMY OF SCIENCES

Volume 30 September 15, 1944 Number 9

Copyright 1944 by the National Academy of Sciences

*GENETIC LINKAGE OF BLOOD TYPES IN THE RABBIT**

BY P. B. SAWIN, M. A. GRIFFITH and C. A. STUART

DEPARTMENT OF GENETICS, BROWN UNIVERSITY

Communicated July 7, 1944

Probably no human characteristics have been subjected, from a genetic standpoint, to more continuous and wide-spread observation, than have blood differences, the inheritance of which is accepted as being typically Mendelian. Yet association in inheritance of a human blood group or serological character with a morphological character has not been demonstrated, except in one doubtful case.[5]

With animals other than man the situation up to the present time has been similar. We are fortunate, therefore, in being able to report a first clearly demonstrated case of genetic linkage in the rabbit between a serological character, the so-called *A* character (presence of the *A* antigen *An* vs. the α agglutinin, *an*),[11, 13,14] and the morphological character brachydactyly (*br*).

We are also able to make a progress report on a study as yet incomplete in which genetic linkage is shown *not* to exist between the serological H_6 antigen and nine different and independent morphological characters. It remains to investigate the linkage relation of H_6 to two other morphological characters, satin (*sa*) and ataxia (*ax*).

The H_6 antigen is one of several antigenic differences known in the rabbit. The serological properties of the H_1 and H_2 antigens have been described by several authors and the inheritance of H_1 and H_2 has been studied and described in a series of publications by Castle and Keeler[2, 3, 7] who found H_1 and H_2 to be alleles and genetically independent of five known genes, viz., absence of the agouti pattern (*a*), dilution of coat color (*d*), non-extension of black pigment (*e*), the dominant white spotting (English) gene (*En*) and the short-haired (rex) gene (r_1). H_1 and H_2 were also found by Knopfmacher[8] to be independent of the following six genes: *a*, *an*, *ax*, *c*, *e* and *sa*. See Knopfmacher's paper for details and page 220 of this paper for significance of gene symbols.

The identity of the H_6 antigen as serologically distinct from and genetically independent of H_1 and H_2 was demonstrated by Knopfmacher[9]. In extension of his work linkage tests have now been made which show the independence of H_6 from the mutant genes *a*, *br*, *c*, *d*, *e*, *En*, r_2, r_3 and *v*. See table 1. Study has not yet been made of the relation of H_6 to genes *sa* and *ax*, so it is still possible that a linkage may exist between H_6 and one of these genes.

TABLE 1

LINKAGE TESTS OF THE H_6 ANTIGEN WITH OTHER KNOWN GENES OF THE RABBIT

GENE	NUMBER OF PROGENY	NON-CROSSOVER	CROSS-OVER	DEVIATION	P.E.	DEV./P.E.
a	37	19	18	0.5	2.05	0.243
c	42	20	22	1.0	2.19	0.456
e	72	36	36	0.0	. .	. . .
d	45	26	19	3.5	2.26	1.558
r_2	95	40	55	7.5	3.29	2.279
En	49	26	23	1.5	2.36	0.635
v	104	47	57	5.0	3.44	1.395
br	118	66	52	7.0	3.66	1.912
r_3	92	44	48	2.0	3.23	0.619

The usual method of linkage analysis, backcross of the double heterozygote to the double recessive, has been employed in these tests using the coupling heterozygous combination (AB/ab) in all cases except in the test involving English spotting. In this case English being a dominant mutation, the repulsion combination is more easily employed. Antiserums for the identification of blood types have been produced by the general method described by Knopfmacher,[9] with the exception that the injections made on successive days were progressively increased by 0.5 ml. until the last injection of 4 ml. was made. The animal was then rested five days and on the sixth day given a final dose of 4 ml., and bled five to seven days later. The antiserums used had a titer of at least 640. The H_1 antiserum had a titer of 1280. In typing animals the procedure described by Knopfmacher was routinely used and tests were made usually at birth for H_1, H_2 and H_6 at the same time.

Some irregularities such as were encountered previously by Castle and Keeler[2] and Battey[1] have also been observed in this study. Animals previously typed as negatives (lacking H_2 and H_6 antigens) during the summer months, when retyped during the cooler months of the year were found to possess these antigens. Some evidence, that such fluctuations are actually influenced by variations in temperature, has been obtained by Griffith[6] in a study of twelve animals possessing H_2 and H_6 antigens at normal room temperatures. When subjected over a period of two weeks, to temperatures ranging from 89–108°F., a definite drop in the strength of the antigen was observed. In some cases it disappeared entirely, but, in all cases, rose to the original titer with the return of the animal to normal room temperatures. To avoid uncertainty the results of the linkage tests of H_6 with other known genes shown in table 1 include only observations made during the period from September 15 to June 15, when room temperatures are regularly lower than those which resulted in diminished titers.

Tests for the first two genes (*a* and *c*) listed in table 1 were completed but not published by Knopfmacher. The others have been made by us, but all alike show no statistically significant departure from the equality of crossover and non-crossover classes expected in independent assortment.

In studying the *A* character (*an*), both agglutination and double complement fixation tests were made for each animal. Since α agglutinins are not present until rabbits are about three months old, neither test was made prior to that age. In cases of uncertainty or disagreement between the results of agglutination and inhibition tests, retests were made, the animal usually being bled a few weeks later.

TABLE 2

LINKAGE TESTS OF THE *A* CHARACTER (*an*) WITH SATIN (*sa*), VIENNA WHITE (*v*) AND BRACHYDACTYL (*br*)

GENE	TYPE OF MATING	NUMBER OF PROGENY	NON-CROSS-OVER	CROSS-OVER	DEVIATION	P.E.	DEV./P.E.	PER CENT CROSSING OVER
sa	Coupling	69	33	35	1.0	2.78	0.350	....
v	Coupling	106	58	48	5.0	3.47	1.440	....
br	Coupling	82	49	33	8.0	3.05	2.622	40.24 ± 3.72
br	Repulsion	32	23	9	7.0	1.91	3.644	28.12 ± 5.96
br	Total (*br*)	114	72	42	15.0	3.60	4.166	36.84 ± 3.16

Adequate tests for association between the *A* character and the mutant genes *a*, *c*, *d* and *En* were made by Sleeper[12] with negative results. Tests subsequently made by us for association between the *A* character (*an*) and the mutant genes *sa*, *v* and *br* are summarized in table 2. In the tests with *sa* and *v*, no evidence was found of linkage. But in relation to the *br* gene, evidence of linkage was found both in the coupling and in the repulsion matings. Of the 82 progeny obtained in the coupling matings, 33 (or 40.2 ± 3.7 per cent) were crossovers. In the repulsion matings 9 of the 32 progeny were crossovers, or 28.1 ± 6.0 per cent. Combining the two experiments gives us a population of 114 backcross individuals, in which 42 are crossovers, or 36.8 ± 3.2 per cent, a deviation from equality (expected if assortment is free) more than four times the probable error and so highly significant. This estimate is based entirely on segregation occurring among the gametes of F_1 females, in which the crossover percentage is regularly higher than in males, as has been demonstrated for rabbits, rats and mice, and is probably true for mammals in general.

It was shown by Castle and Sawin[4] that brachydactyl (*br*) is linked with furless (*f*) with a crossover percentage of about 28.3. Since this is considerably less than the estimated 36.8 per cent of crossovers between *br* and the *A* character (*an*), it seems probable that the order of the three genes is *br–f–an* rather than *f–br–an*. The former gives a total map distance of less than 40, whereas the latter would make it in excess of 65, an improbably high figure.

Summary.—It has long been known that serological differences, such as the blood groups of man and other mammals, are inherited in typical Mendelian fashion and so are undoubtedly gene controlled. The first clear evidence of linkage between a serological gene and a gene for a morphological character, brachydactyl, is presented herewith. A third gene, furless, lies in the same linkage system with these two in what has been called the fifth chromosome of the rabbit.

Bringing up to date our knowledge of linkage in the rabbit, the six demonstrated linkage systems may be diagrammed thus:

Chromosome				Chromosome			
	c	*y*	*b*	IV	*a*	*dw*	*w*
I							
	0	14.4	42.8		0	14.7	29.9
	du	*En*	*l*		*br*	*f*	*an*
II				V			
	0	1.2	14.3		0	28.3	36.8
	r_1	r_2			*E*		*At*
III				VI			
	0	17.2			0		26.2

GENE SYMBOLS

a,	non-agouti	*dw*,	dwarf
an,	antigen *A*	*e*,	non-extension (yellow)
at,	atropinesterase[10]	*En*,	English spotting
ax,	ataxia	*f*,	furless
b,	brown (chocolate)	r_1,	rex No. 1, short hair
br,	brachydactyl	r_2,	rex No. 2, short hair
c,	albino	r_3,	rex No. 3, short hair
d,	dilute (blue)	*w*,	wide band
du,	Dutch spotting	*y*,	yellow fat

---

* A portion of a research program supported in part by a grant from the Rockefeller Foundation.

[1] Battey, S. O. M., Thesis, Brown University (1940).

[2] Castle, W. E., and Keeler, C. E., *Proc. Nat. Acad. Sci.*, **19**, 92 (1933).

[3] Castle, W. E., and Keeler, C. E., *Ibid.*, **19**, 98 (1933).

[4] Castle, W. E., and Sawin, P. B., *Ibid.*, **27**, 519 (1941).

[5] Finney, D. J., *Ann. Eugenics*, **10**, 171 (1940).

[6] Griffith, M. A., Thesis, Brown University (1944).

[7] Keeler, C. E., and Castle, W. E., *Proc. Nat. Acad. Sci.*, **19**, 403 (1933).

[8] Knopfmacher, H. P., Thesis, Brown University (1942).

[9] Knopfmacher, H. P., *J. Immunol.*, **44**, 121 (1942).

[10] Sawin, P. B., and Glick, D., *Proc. Nat. Acad. Sci.*, **29**, 55 (1943).

[11] Sawin, P. B., Stuart, C. A., and Wheeler, K. M., *J. Heredity*, **34,** 179 (1943).

[12] Sleeper, H. A., Thesis, Brown University (1940).

[13] Stuart, C. A., Sawin, P. B., Wheeler, K. M., and Battey, S. O. M., *J. Immunol.*, **31,** 31 (1936).

[14] Wheeler, K. M., Sawin, P. B., and Stuart, C. A., *Ibid.*, **36,** 349 (1939).

CONCERNING LINKAGE OF WALTZING IN RATS

By MAURICE WHITTINGHILL

DEPARTMENT OF ZOÖLOGY, UNIVERSITY OF NORTH CAROLINA

Communicated July 24, 1944

Some mistakes in recent papers[1] on linkage in rats should be pointed out not only for the sake of accuracy in rat genetics but because the errors are of types made all too frequently in other fields. The mistakes stem from experiments that approach the problem repeatedly from one angle, which may result in an incompletely supported conclusion. They also include correct verbal logic inaccurately related to the experimental events and misapplication of the formula for computing the probable error to data which had been adjusted because of experimental necessity.

The idea that the recessive factors waltzing and albinism are linked seems to be a premature conclusion from experiments revealing only one side of the picture. In all the crosses reported in both papers the two genes were introduced into the F_1 through the same parent and were found together oftener than not in the F_2 and in testcross offspring. A variety of experiments of this coupling type, involving over 1000 rats, gave consistent results; yet there were no experiments to show whether the two factors, if introduced into the F_1 from *opposite* parents or in repulsion phase, would tend to remain *separated* in the gametes of these heterozygotes. Without such information from repulsion matings it is not evident how one can distinguish linkage of the two genes in inheritance from embryological enhancement of the waltzing in the albino phenotype. The latter is more than a remote hypothesis, because waltzers overlap with wild and are recognizable much less frequently than expected on the single factor hypothesis.

A second mistake, while made apart from the first, contributes further to the confusion between events in inheritance and in development. It is known that the recessive pink-eyed yellow gene is located in the same chromosome *pair* as the albino allels. Consequently, King and Castle say (p. 57), "If waltzing is linked with albinism, it should show linkage also with yellow." They performed a three-point experiment in which, it should be carefully noted, waltzing and albinism were introduced into the F_1 from one parent while yellow was inherited from the other. Assuming linkage, one would expect from such a heterozygote an excess of those gametes containing either one or the other of the original combinations: waltzing, albinism, non-yellow or else non-waltzing, colored, yellow. In other words the waltzer and albino genes would most of the time remain coupled with each other and repulsed from yellow during meiosis. In linkage inheritance certain genes tend to be held together in one chromosome while their allels are apt to be coupled in another group which is repulsed from the first. Forgetting that yellow would show repulsion from albinism and waltzer according to their hypothesis, the authors state (p. 58), "As expected, the percentage of waltzers in the yellow group is higher than that in the non-yellow, suggesting linkage of waltzing with yellow as well as with albinism."

TABLE 1

OFFSPRING OF CASTLE AND KING'S TESTCROSS

$$F_1 \frac{C\ p\ W}{c\ P\ w} \text{ BY } \frac{c\ p\ w}{c\ p\ w}$$

GAMETES	CROSSOVERS IN REGION	OFFSPRING RECORDED (ADJUSTED NUMBERS IN PARENTHESES)			
		GRAY, $Cc\ Pp$	YELLOW, $Cc\ Pp$	ALBINO, $cc\ ?\ p$	
$C\ p\ W$	0	......	180 (129.6) normals	....	
$c\ P\ w$	0	......	....	S 66 (145.2) waltzers in two S classes	
$C\ P\ w$	1	11 (24.2) waltzers	....	....	
$c\ p\ W$	1	......	....	T 175 (95.8) normals, in two T classes	
$C\ p\ w$	2	......	42 (92.4) waltzers	....	
$c\ P\ W$	2	......	....	T	
$C\ P\ W$	1 and 2	49 (35.8) normals	....	....	
$c\ p\ w$	1 and 2	......	....	S (see $c\ P\ w$)	
Totals		60	222	241	523

Let us arrange the data from Castle and King's three-point testcross in a table to show (1) the occurrence of single, double and non-crossovers on the linkage hypothesis, and (2) the proportion of waltzers in each of the three coat-color classes. I am grateful to Dr. Castle for a suggestion which helped to simplify this table. Inspection of table 1 does not reveal the usual result of crossing-over, namely, a small number of double crossovers as compared with each of the single crossover classes. Since the color classes may be under suspicion of differentially enhancing waltzing, let us consider them separately. Among the gray rats there seem to be too many double exchanges. Among the albinos too few non-crossovers appear. Among the yellows the percentage of crossovers between the *p* and *w* loci is low in comparison with the total data. In all cases the low classes are waltzer classes, which comprise only 22.8% plus or minus the standard error of 1.76%, whereas one would expect 50% waltzers if waltzing is differentiated from the normal by a single pair of genes. Consequently Castle has proposed a reasonable correction for this lack of waltzers.[2] He has added 1.2 times the waltzers obtained to each waltzer class at the expense of the corresponding non-waltzers. This gives the adjusted numbers shown in parentheses in table 1, where waltzers now equal non-waltzers in numbers. Other adjustments do not seem to be in order, since the 282 coloreds (grays plus yellows) are about equal to the 241 albinos; and the ratio of grays to yellows is an expression of the amount of crossing-over between those two loci, 22 units apart as shown here and previously.

TABLE 2

RECOMBINATION BETWEEN THE YELLOW AND WALTZER LOCI (I.E., ASSUMED CROSSOVERS IN REGION 2)

	%	S.E., %
Region 2 crossovers among grays (raw data)	81.7	± 4.96
Region 2 crossovers among grays (adjusted)	59.7	± 4.96
Region 2 crossovers among yellows (adjusted)	41.6	± 2.63
Region 2 crossovers among yellows (raw data)	18.9	± 2.63
Combined adjusted data	45.5	± 2.96
Probability of such a chance deviation from a 50% recombination, 0.13		

The effect of adjusting the data in the manner described is, among the coloreds, entirely upon the value obtained for region 2 (see table 2). In this table it will be noticed that the standard errors attached to the adjusted percentages of crossovers are the same as those computed for the actual samples. One could obtain a larger sampling error by treating the adjusted percentage as a raw sample, but I have not done so. The difference in the raw data is obviously great. The difference between the crossover values given by the gray classes and by the yellow classes, after the data have been adjusted in respect to waltzing, is still a very significant difference indicating that something more than linkage is operating. If the values of 59.7 and 41.6% crossovers had come from raw samples rather than from the adjusted figures nevertheless the difference between them could have occurred by chance alone only 1% of the time and should therefore be adjudged a real difference and one not to be merged by averaging. Considered another way, the 41.6% of yellows has a 0.016 probability of being a chance deviation from a 50:50 distribution; and the 59.7% of crossovers among grays could readily be a chance departure from independent assortment, *P* being 0.13, but in the opposite direction. Considered together the probabilities are thus better for independence than for linkage. Even if these differing measures of recombination are averaged, the result is on the borderline of a significant departure from independent assortment.

Using the raw data an important comparison is possible among those rats of this testcross which actually show waltzing. This comparison should be very meaningful because no adjustment, however reasonable, has been made for the expression of waltzing. All colored waltzers result from recombination, either in region 1 or in the provisional region 2. About 18% of each color are waltzers. The colored waltzers are distributed as follows: 11 gray waltzers, which would arise only from crossing-over in region 1; and 42 yellow waltzers, which result from an exchange in region 2. It follows from this that region 2 must be several times longer than the 22-unit distance contained in region 1. That would mean an exchange frequency great enough to represent free assortment between pink-eyed yellow and waltzing.

Since the correction for the poor expression of waltzing does not establish linkage to the satisfaction of this author, let us now see if any constructive conclusions using the contrary hypothesis of independent assortment may yet be drawn as to the penetrance of waltzing in different color classes. Among gray rats, which result from crossing-over between the *c* and *p* loci, waltzing is expressed in 18.3 ± 4.96% of the rats instead of the 50% expected without linkage. This gives a penetrance of 36.6% meaning that only 36.6% of the genetically waltzer rats showed that character phenotypically. Among yellow rats the proportion of waltzers is the same, 18.9 ± 2.63%. Assuming independent assortment of *w* and *y*, we would expect 50% waltzers, so the penetrance here is 37.8%. If linkage should be involved, the expected percentage would be below 50, and the resulting penetrance, for various assumed strengths of linkage, is shown in the right hand side of table 3.

In the remaining class, the albinos, the occurrence of waltzing is 27.5 ± 3.4%, higher than and perhaps significantly different from the frequency of waltzers among the grays and the yellows. The differences are twice the standard error of the respective differences, so that the odds are about 1 against 20 for them to be merely chance deviations. In the albinos the penetrance of waltzer turns out to be 55% if we assume no linkage. On the other hand, if we say that the penetrance is the same as in the colored classes, close linkage with only 10 or 20% of crossing-over would be indicated. No such close linkage has been found. Table 3 shows these and intermediate correspondences between crossing-over and enhancement of expression. The percentage of waltzers expected takes into account the merging of the double crossovers phenotypically into the non-crossover class as shown in table 1. It is assumed that double crossing-over takes place without interference, so that the frequency of double (equals the product of all exchanges for each of the two regions figured separately.

It may be true that both augmentation and linkage are responsible for this situation. In that case some other value of crossing-over, between the 35% tentatively arrived at by Castle and King (1941) and the free assortment suggested by myself, may turn out to be correct. A direct measure of penetrance could be obtained through data from inbred albino waltzer stock or from yellow waltzer stock independent of linkage complications. Separate measures of penetrance might through table 3 confirm some particular value for crossing-over. Perhaps the difficulty in establishing the desired yellow waltzer stock is a low penetrance of waltzing in yellow rats. Within the limits of sampling errors the penetrances of waltzer in albinos and in yellows should lead to the same conclusion as to the frequency of recombination in the long and doubtful region. A more direct and unequivocal test would be an experiment starting with yellow waltzer mated to albino and testcrossing the F_1, but technical difficulties often slow one's realization of the conclusive demonstration.

TABLE 3

DEGREES OF PENETRANCE FOR THE WALTZER GENE, *w*, AS FIGURED FROM CASTLE'S TESTCROSS IF LINKAGE IS ACTING

ASSUMED % OF CROSSOVERS BETWEEN YELLOW AND WALTZER	% WALTZERS AMONG ALBINOS		PENETRANCE OF *w* IN ALBINOS, %	% WALTZERS AMONG YELLOWS		PENETRANCE OF *w* IN YELLOWS, %
	EXPECTED	FOUND		EXPECTED	FOUND	
10	72	27.5	38	10	18.9	..
20	67	27.5	41	20	18.9	95
30	62	27.5	44	30	18.9	63
40	56	27.5	49	40	18.9	47
45	53	27.5	51	45	18.9	42
Free assortment	50	27.5	55	50	18.9	38

A third type of mistake is a misapplication of a formula for computing the sampling error to data which do not comprise an actual sample. In the testcross transcribed in table 1 it was not possible to distinguish the albinos from the albino yellows, hence one of the two complementary crossover classes for region 1 could not be recognized. Sixty of the other class were found in a total of 523 offspring of all kinds. It was reasonable to double this 60 before dividing by the total in order to get the most likely per cent of recombination, visible and concealed. In figuring the reliability of this sample, the total of 523 seems to have been used giving a probable error of 1.2%. It would be more appropriate to use only the actually observed number of crossovers, 60 grays, and to divide that by the number of non-albinos, or 282 coloreds. By thus removing the unclassifiable albinos a greater probable error is revealed, 1.6%. Even without this leeway the new crossover value for region 1, 21.3%, agrees better with the previous result of 21.9 ± 0.4% obtained by Castle and Wachter.

---

[1] King, Helen Dean, and Castle, W. E., these PROCEEDINGS, **23**, 56–60 (1937); Castle and King, *Ibid.*, **27**, 394–398 (1941).

[2] Castle and King, *Ibid.*, **27**, 396 (1941).

LINKAGE OF WALTZING IN THE RAT

By W. E. Castle

Division of Genetics, University of California

Communicated July 24, 1944

I am greatly indebted to Dr. Whittinghill for pointing out in his table 1 of the foregoing paper a serious mistake which I made in calculating crossover percentages in the data for the three-point cross involving waltzing, albinism and pink-eyed yellow (Castle and King, 1941). The mistake consisted in not making allowance for *double crossing-over*, i.e., crossing-over occurring simultaneously in regions *C P* and *P W* of the albino chromosome.

While the occurrence of linkage as originally assumed by Castle and King would still be shown, the map distances would be materially altered.

Let us then reëxamine the entire evidence in which linkage (coupling) between albinism and waltzing was postulated in 1937 and reasserted in 1941 with the additional conclusion that linkage in the repulsion phase was found between pink-eyed yellow and waltzing.

In the earlier (1937) paper cited, an attempt was made to summarize and interpret the results of three different crosses made between albino waltzers and colored non-waltzers known as hairless, kinky and blue, respectively.

All F_1 individuals were colored non-waltzers showing the recessive nature of both albinism and waltzing. In each of the F_2 populations four phenotypes were obtained, viz., colored non-waltzers (*C W*), albino non-waltzers (*c W*), colored wlatzers (*C w*), and albino waltzers (*c w*). In each of the F_2 populations fewer waltzers were obtained than were to be expected if waltzing is a simple recessive. This, however, is in accord with numerous earlier observations of King and is to be explained by the low effectiveness (penetrance) of the waltzing gene, so that many individuals which are homozygous for waltzing nevertheless do not waltz. The percentage of individuals which actually waltz is far below the expected 25% but higher among F_2 albinos (12%) than among F_2 colored individuals (5%), which suggests the occurrence of linkage (coupling) between the recessive genes *c* and *w*, which entered the cross together.

To test this suggestion backcrosses were made between double recessive individuals (albino waltzers) and F_1 colored non-waltzers. From such matings 1:1 ratios were to be expected of colored and albino individuals, as well as of waltzers and non-waltzers, if both albinism and waltzing were single gene characters with normal complete penetrance. Substantial equality was realized between albino and colored classes, viz., *1318* albinos:*1365* colored. But again the incidence of waltzing was higher among the albinos (29%) than among the colored (22.7%), whereas 50% was to be expected in each case if penetrance were complete and linkage absent. Dr. Whittinghill has suggested that waltzing may be easier of realization in an albino than in a colored phenotype and that this rather than linkage may be the explanation of the difference in percentage of waltzers among albinos and among colored individuals. The suggestion is negatived by the result of a backcross in which a black waltzer (heterozygous for albinism) was mated with F_1 black individuals (cross 5 *C*, p. 58). Here the percentage of waltzers in the backcross population (20.9) was about the same as in other backcross populations, but a majority of the waltzers (43) were *black* while only 9 were albinos. This shows that waltzing is realizable in the usual percentage of cases in a colored as well as in an albino phenotype, and that the phenomenon under discussion is genuinely genetic, not merely developmental.

Dr. Whittinghill's further contention that a repulsion cross in addition to a coupling cross would be required for the complete demonstration of genetic linkage may be dismissed in the light of the repeatedly demonstrated fact that both in plants and in animals linkage in substantially like amounts will be shown by coupling and repulsion crosses involving in both cases the same pair of genes. The desirability of such supporting evidence may be freely admitted, but practical considerations at the time made it difficult to obtain.

We come now to a second line of evidence that linkage exists between albinism and waltzing, evidence which also is called in question by Dr. Whittinghill. Extensive experiments made many years ago showed that albinism (*c*) and pink-eyed yellow (*p*) are linked with about 22 per cent of crossing-over between them. Consequently, if waltzing is linked with albinism, it should show linkage also with yellow. The experiments described in our 1937 paper did not to our knowledge involve the yellow gene, until this fact was revealed by the appearance of the yellow phenotype in F_2 both in waltzing and in non-waltzing individuals. Then it became evident that the yellow gene had been introduced in our cross 3 as a recessive in the blue (colored, non-waltzing) parent. Among 63 F_2 yellows, 6 were waltzers. The number was too small to have statistical significance, but in percentage frequency (9.5) was intermediate between that observed in other colored phenotypes (5.8) and that observed among the albinos (14.7). Dr. Whittinghill has quite correctly pointed out that "forgetting" the repulsion relationship in which the yellow gene had been introduced we hazarded an unsupported guess that waltzing might be linked with yellow as well as with albinism. Of course there was no statistically significant evidence for or against the suggestion and we did not assert that there was. But we did say (p. 58) that "Further experiments to test for linkage between waltzing and yellow [already in progress] are as yet incomplete and inconclusive." To obtain an adequate amount of data from such experiments required several years additional work. The results are contained in our 1941 paper.

A cross was made between pink-eyed yellow non-waltzers and albino waltzers. The F_1 animals were gray non-waltzers and their genetic constitution was reported thus, $\frac{C\ p\ W}{c\ P\ w}$. We were perfectly aware that this time p had been introduced in the repulsion relationship to waltzing for we had deliberately planned it that way.

After triple recessives ($cc\ pp\ ww$) had been laboriously synthesized, which in every gamete would transmit the three recessive allels, $c\ p\ w$, Dr. King backcrossed such triple recessives to F_1 mates obtaining a population of 523 animals distributed as follows:

	NORMAL	WALTZING	TOTAL	% WALTZERS	% CROSSOVER
Gray	49	11	60	18.3	18.8 ± 1.5 (C/w repulsion)
Yellow	180	42	222	18.9	
Albino	175	66	241	27.4	72.6 ± 1.8 ($c\ w$ coupling)
Total	404	119	523	20.8	

That albinism is linked (coupled) with waltzing is indicated by the high percentage of albino waltzers (27.4). That the color gene (C) and the yellow gene (p) are *repelled* by waltzing is indicated by the much lower percentage of waltzers in the gray and the yellow classes (18.3 and 18.9, respectively).

Consider first the detailed evidence concerning the color gene (C) in relation to waltzing (w). All colored individuals would have derived C from their F_1 parent, which would produce gametes either $C\ W$ (colored non-waltzers, gray or yellow) or $C\ w$ (colored waltzers, gray or yellow). The progeny were 49 gray + 180 yellow = 229 non-waltzers, and 11 gray + 42 yellow = 53 waltzers, an indicated crossover percentage (repulsion between C and w) of 18.8 ± 1.5. From the albino group we may obtain additional evidence as to the linkage relations between C (or rather its recessive allele, c) and w. Gametes produced by the F_1 parent which would give rise to albino offspring would be either $c\ W$ (crossovers, non-waltzers) or $c\ w$ (non-crossovers, waltzer). The numbers recorded are 175 crossovers to 66 non-crossovers, an indicated but highly improbable crossover percentage of 72.6. Such are the results based on the raw data the p locus being ignored, as it was in our 1937 experiments, in which the p gene was not involved except in the single experiment designated cross 3, as shown by the F_2 population of table 3, page 58.

Obviously it is necessary in dealing with a character having such low penetrance as waltzing manifests, to make a correction for the deficiency of phenotypic waltzers, as we did in our 1937 paper. This can be done by increasing the waltzer classes to their theoretical equality with half the entire population, and decreasing by like amounts the corresponding non-waltzer classes. In the present case, this may be accomplished by multiplying the recorded numbers of waltzers by the factor 2.2. The frequencies in each group then become as follows:

	NON-WALTZERS	WALTZERS	TOTAL	% WALTZERS	% CROSSOVER
Gray	35.8	24.2	60	40.3	41.3 ± 1.9 (C/w repulsion)
Yellow	129.6	92.4	222	41.6	
Albino	95.8	145.2	241	60.2	39.8 ± 2.1 ($c\ w$ coupling)
Total	261.2	261.8	523	50.0	

Waltzers now, it will be observed, still appear as minority groups among the gray and the yellow classes, owing to repulsion between C and w; but among the albinos waltzers become a majority group, owing to coupling between c and w. The crossover percentage for repulsion in the combined gray and yellow classes (the entire colored population) is 41.3 ± 1.9, for coupling in the albino class it is 39.8 ± 2.1. For the entire population (colored and albino, repulsion and coupling experiments combined) it is 40.6 ± 1.4.

The value obtained for $C\ w$ in our 1937 paper in which locus p was not involved was 44.7 ± 0.7 based on a population of 2139 backcross individuals. When a correction was made by increasing the number of waltzers to half the total population, the figure became 45.8 ± 0.7. It seems probable therefore that the true value lies between 40 and 45.

But if the p locus is taken into consideration, as it should be in a three-point cross, it is obvious, as Dr. Whittinghill points out that double crossing-over must be reckoned with, i.e., crossing-over simultaneously in region 1 ($C\ p$) and in region 2 ($p\ W$) which would result in the production of triple dominant gametes ($C\ P\ W$) potentially gray non-waltzers. These number 49 in the raw data and 35.8 in the corrected data.

Using first the raw data, the indicated crossovers in region 1 are 11 gray waltzers + 49 gray non-waltzers = 60 individuals or 21.3% of the colored population. The indicated crossovers in region 2 are 42 yellow waltzers + 49 gray non-waltzers = 91 individuals, or 32.2 ± 1.8% of the colored population.

On this basis the linkage map for the $C\,p\,W$ portion of the albino chromosome would be $\frac{C\ 21.3\ P\ 32.2\ W}{0 \qquad\qquad 53.5}$. The finding 21.3 for the $C\,P$ region is in good agreement with the value 21.9 ± 0.4 found by Castle and Wachter (1924).

Let us now see how these values are affected by using the corrected data. The total amount of crossing-over indicated in region 1 (region $C\,p$) remains unchanged at 60 (24.2 gray waltzers + 35.8 gray non-waltzers) or again 21.3% of the colored population, which is in complete agreement with extensive earlier investigations. In region 2 ($p\ W$) the indicated crossovers are 92.4 yellow waltzers + 35.8 gray non-waltzers = 128.2 or 45.4 ± 1.97% of the colored population. Dev./P.E. = 2.3. A deviation from 50% of like magnitude might be expected by chance 12 times in 100, the odds against it being 7.28 to 1.

The linkage map would now become $\frac{C\ 21.3\ P\ 45.4\ W}{0 \qquad\qquad 66.7}$. The $C\,W$ interval calculated without regard to p and based on the colored population only was 41.3 ± 1.9. When p is taken into consideration and double crossing-over allowed for, the interval becomes 53.5 based on the raw data, 66.7 based on the corrected data. Repulsion between p and w is 32.2 ± 1.8% based on the raw data, 45.4 ± 1.97% based on the corrected data. Calculated either way statistically significant coupling is shown between waltzing and albinism, and statistically significant repulsion between waltzing and the pink-eyed yellow gene is shown to be highly probable.

In the light of these findings the linkage map of the albino chromosome must be revised to show it considerably more extended than in our 1941 paper. It would appear about like this:

$$\frac{L\quad C\quad R\quad P\quad W}{0\ \ 3.3\ \ 3.8\ \ 25.2\ \ 70.6}.$$

Dr. Whittinghill's suggestion that it is desirable to base the probable error on the half-population under the circumstances described by him is entirely correct, as is also his insistance that double crossing-over and interference are not synonymous terms (p. 396). An error which he charitably calls "typographical" was due to an unfortunate aberration when I wrote non-curly instead of non-anemic in the second part of the 1941 paper (p. 398, line 9). I am obliged to him for pointing out these errors.

The slight differences between Dr. Whittinghill's figures and my own in certain instances are due to the fact that he uses standard errors and I probable errors.

X-RAY INHIBITION OF MITOSIS IN RELATION TO CHROMOSOME NUMBER*

BY A. MARSHAK AND MURIEL BRADLEY

RADIATION LABORATORY AND DEPARTMENT OF BOTANY, UNIVERSITY OF CALIFORNIA

Communicated July 12, 1944

Numerous references to mitotic inhibition resulting from ionizing radiations appear in the literature, but the problem of the mechanism of this inhibition is still unsolved. Strangeways and Oakley,[1] Strangeways and Hopwood[2] and several subsequent investigators have shown that cells already in mitosis at the time of irradiation continue to divide, but that cells which normally would have entered mitosis during that period are inhibited at onset of prophase. Evidence that the inhibition is caused by the action of x-rays upon the nucleus was supplied by Henshaw and Francis (1936)[3] and by Henshaw (1938).[4] They found that x-rayed *Arbacia* sperm used to fertilize untreated eggs will delay cleavage, and Henshaw has shown that whereas irradiation of enucleated egg fragments has no effect on time of division irradiation of nucleated fragments caused delay in division. Crowther[5] suggested that the centrosome might be the structure involved in the inhibition, since the centrosome and the sensitive volume (calculated from the data of Strangeways and Oakley[1]) were of the same order of magnitude. Fröier, Gelin and Gustafsson[6] concluded that extra-chromosomal processes within the nucleus influence the onset of prophase and, when disorganized by x-rays, they prevent mitosis. This inference was based on the ability of cells to divide even when very high x-ray doses have broken up the chromosomes to the extent that they have lost all distinct continuity and shape. However, other interpretations will fit their observations. Previous investigation by one of us showed that there is no correlation between chromosome length and mitotic inhibition.[7] In this paper we wish to report evidence that inhibition of mitosis is dependent upon the number of chromosomes in the cell.

Root tips of three species of *Triticum* (*T. monococcum*, *T. dicoccum* and *T. vulgare*) and two species of *Bromus* (*B. carinatus* and *B. arizonicus*) were selected as the biological material. Both genera are members of the Gramineae and both have basic diploid chromosome numbers of 14. The numbers of chromosomes in these five species form a series of 14, 28, 42, 56 and 84. The source of x-rays was an oil-cooled tube with a tungsten target in a General Electric Company Maximar apparatus, operated at 220 Kvp. and 15 ma. The beam was filtered through 1 mm. Al, and a lead diaphragm 1 cm. thick and with a 10 × 10 cm. aperture was used at the base of the cone which was 17.5 cm. in length. The target distance was 85 cm. Irradiation was given at the rate of 43 r per minute as measured with a Victoreen ionization chamber.

Seeds were germinated on moist filter paper in covered glass dishes kept at a constant temperature of 25°C. When the roots were 2 to 3 cm. long, seedlings of each species were irradiated in five different lots with increased dosage for successive lots. After irradiation the seedlings were returned to the incubator for 3 hours. Cells at the onset of prophase when irradiated are in anaphase 3 hours later.[7] At the end of that interval root tips were fixed in a solution of 1 part 50% ethyl alcohol : 1 part 50% acetic acid. The fixation period was 18 to 20 hours, after which the root tips were left in 70% alcohol for approximately 24 hours. They were then washed in distilled water for 10 minutes, treated with 4% iron alum for 1 hour, washed in distilled water for 1 hour, and macerated in concentrated hydrochloric acid for 1$^{1}/_{2}$ minutes. After washing them again in distilled water for 20 minutes or longer, aceto-carmine smears were made. Sufficient length of each root tip to insure the inclusion of the entire region of cell division was smeared. Two small drops of aceto-carmine were placed close together on the slide, in one of which two root tips were gently

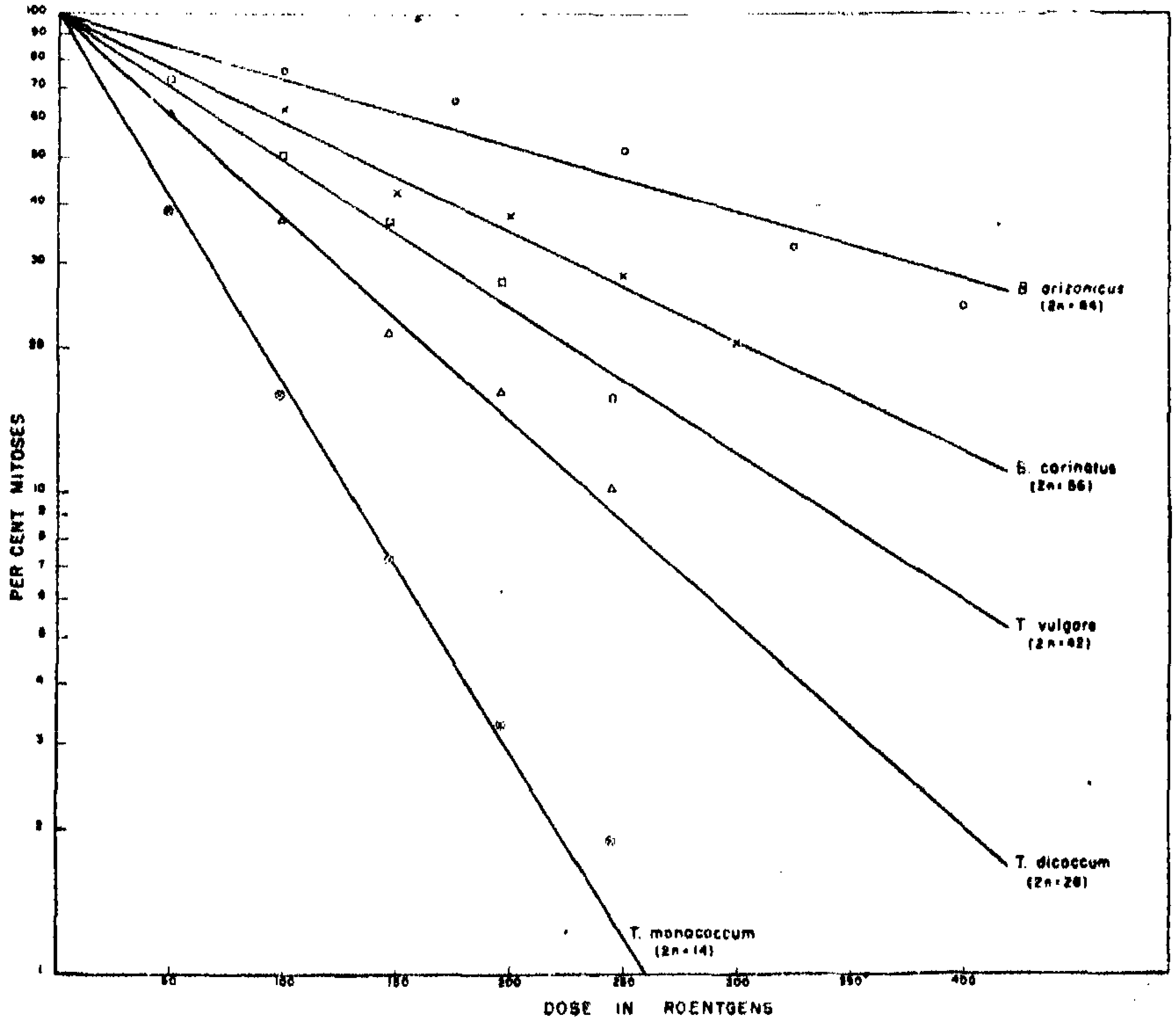


FIGURE 1

Logarithms of per cent mitoses as a function of dose in roentgen units.

mashed with a dissecting needle to separate the cells. Cells adhering to the needle were then washed off in the other drop of stain and a cover glass placed over both drops and their contents. To prevent the escape of any cells from under the cover glass, care was taken to use just enough aceto-carmine to extend to the edge of the cover glass when the smear was adequately flattened. Control root tips for each experiment received treatment identical with that given irradiated root tips, with the exception of exposure to x-rays.

All metaphases and anaphases were counted in 8 to 18 root tips of each control or irradiated lot. Prophases and telophases were not included in the counts because of difficulty in distinguishing accurately between late resting stage and early prophase and between late telophase and early resting stage. The maximum variation between mitotic counts for any two slides of one lot are as follows: *T. monococcum* 11%, *T. dicoccum* 34%, *T. vulgare* 40%, *B. carinatus* 25% and *B. arizonicus* 33%. The mean deviations for the same species were 8, 13, 20, 15 and 16%, respectively. Percentages of the average number of mitoses per control root tip were found for each irradiated lot. In figure 1 the logarithms of these percentages are plotted against dosage. The curves are all of the form $Y = e^{-kx}$, where Y = per cent mitoses, e = the base of natural logarithms, x = dose in roentgens and k = a constant. They were fitted by the method of least squares with each point weighted by the number of mitosis observed. In table 1 the slopes of the curves are given, together with the root mean square deviations and the measures of goodness of fit according to the equation

$$R^2 = 1 - \left[\frac{\Sigma(y-Y)^2}{n\left(\frac{\Sigma(y^2)}{n} - \frac{(\Sigma y)^2}{n}\right)}\right]$$

where y = observed log precentage mitoses, Y = the calculated log percentage mitoses and n = the number of points on the curve (Yule and Kendall).[8] Values of the slopes (**k**) and root mean square deviations (S) are given in terms of the base e of natural logarithms.

TABLE 1

GENUS AND SPECIES	CHROMOSOME NUMBER ($2n$)	$k_e \times 10^{-3}$	$S_e \times 10^{-3}$	R
Triticum monococcum	14	17.698	1.29	0.994
Triticum dicoccum	28	9.721	0.53	0.999
Triticum vulgare	42	6.997	0.46	0.999
Bromus carinatus	56	5.197	0.49	0.999
Bromus arizonicus	84	3.159	0.52	0.996

In order to verify previous evidence that inhibition of mitosis is not a function of chromosome length, chromosomes of each species were measured. Two types of measurements were made; in one, all chromosomes of a single metaphase cell from each species were measured from camera lucida projections, with the usual precautions taken to avoid optical distortion. Where a chromosome was not entirely horizontal, its length was determined geometrically from the highest and lowest focal points of those portions which deviated from the horizontal. Since this procedure is very laborious, additional measurements were made with the ocular micrometer. All chromosomes in 10 cells of *T. monococcum*, in 8 cells of *T. dicoccum*, in 5 of *T. vulgare*, in 4 of *B. carinatus* and in 2 cells of *B. arizonicus* were measured by the latter method. Camera lucida and ocular micrometer measurements are listed separately in table 2. Lengths are given in terms of total chromosome length, i.e., the sum of the lengths of all chromosomes in a metaphase configuration.

TABLE 2

TOTAL CHROMOSOME LENGTHS (IN MICRONS)

SPECIES	LENGTH (1 CELL) (CAMERA LUCIDA)	MEAN LENGTH (OCULAR MICROMETER)	RANGE OF LENGTHS (OCULAR MICROMETER)
T. monococcum	124	118 (10 cells)	114 to 127
T. dicoccum	219	203 (8 cells)	194 to 222
T. vulgare	243	248 (5 cells)	238 to 286
B. carinatus	332	287 (4 cells)	274 to 336
B. arizonicus	411	374 (2 cells)	358 to 387

In connection with these data on chromosome length, certain results of the investigations of Fröier, Gelin and Gustafsson[6] are of interest. They have reported the number of cells showing abnormalities following irradiation of dry seeds of the same *Triticum* species used in this work. They point out that there are fewer abnormalities in *T. vulgare* than expected from calculations based on the rate of production of abnormalities in *T. monococcum*. However, their calculations are based on the assumption that the mean length of the chromosomes is the same in the two species. Our measurements show that the total chromosome length in *T. vulgare* is not three times that of *T. monococcum* as assumed by those authors, but is only twice as great. Their observation is in agreement with evidence from a variety of other species that chromosome sensitivity is directly proportional to the total length of the chromonemata or roughly proportional to total chromosome length. Furthermore, from the combined evidence concerning chromosome abnormalities and chromosome lengths one can conclude that the chromosomes of *T. vulgare* are shorter than those of *T. monococcum* because their chromonemata are shorter in absolute length and are not merely more tightly coiled than those of *T. monococcum*.

In table 3 are given the ratios of the chromosome numbers of the different species, the inverse ratios of the slopes of their "survival" curves and the ratios of the total lengths of the chromosomes of their complements.

There is a linear proportionality between the ratios of the chromosome numbers and ratios of slopes, the greatest deviation being 18%. This relationship may be demonstrated in another way: when the slopes are multiplied by the multiples of 14 represented in their respective chromosome numbers, i.e., 1, 2, 3, 4, and 6, the resulting figures are 17.7, 19.4, 21.0, 20.8 and 19.0, or very nearly a constant. No such relationship exists between ratios of chromosome lengths and ratios of slopes.

TABLE 3

RATIOS OF CHROMOSOME NUMBERS	INVERSE RATIOS OF k	RATIOS OF CHROMOSOME LENGTHS (CAMERA LUCIDA)	(OCULAR MICROMETER)
56/42 = 1.33	1.35	1.37	1.16
42/28 = 1.50	1.39	1.11	1.22
84/56 = 1.50	1.63	1.24	1.30
28/14 = 2.00	1.82	1.77	1.72
56/28 = 2.00	1.87	1.52	1.41
84/42 = 2.00	2.21	1.69	1.51
42/14 = 3.00	2.54	1.96	2.10
84/28 = 3.00	3.07	1.87	1.84
56/14 = 4.00	3.42	2.68	2.43
84/14 = 6.00	5.60	3.33	3.18

From these data it becomes evident that for the species studied mitotic survival is directly proportional to multiples of the genome, i.e., with every duplication of chromosome number the survival rate is doubled, or conversely, the mitotic inhibition is reduced by one-half. Since inhibition varies inversely as the chromosome number, but is not similarly related to chromosome length, structures within the chromosomes, either single or if multiple then not at all uniformly distributed along the chromosomes, must be involved in the mechanism of mitosis. Mitotic failure is then caused by inactivation of a minimum number of these structures. The only visible chromosomal structures known which fulfil these requirements are the centromeres (spindle-fibre attachment regions). These results might also be obtained if a small number of genes were involved, but not by a considerable number of genes distributed throughout the length of the chromosomes.

Of these two alternatives, the first is supported by a qualitative result obtained consistently with irradiated cells. Lagging chromosomes, groups of chromosomes or fragments of chromosomes form their own little nuclei when the major portion of the complement enters the telophase. These micronuclei retain their structural integrity until the major nucleus enters the prophase. At that time the micronucleus will go through the characteristic prophase-anaphase changes if its chromosomal components contain centromeres. If they do not, the micronucleus becomes pycnotic and then disintegrates.[9] The presence of the centromere thus appears to be essential for the initiation of those changes associated with the onset of prophase.

If calculations of sensitive volumes are made from the above data, the result obtained is peculiar in that the sensitive volume decreases as the volume of chromosomes is increased. This contrasts strangely with the observations on the production of chromosome abnormalities with x-rays and neutrons, where the sensitivity is directly proportional to the length of the chromonemata and the diameter of the sensitive portion of the chromonema is constant.[7, 10] It differs also from the result obtained with virus and bacteriophage where the calculated sensitive volumes were found to be proportional to the size of the organism as determined by their ability to pass through filters of different pore size.[11, 12] Superficially the results obtained on inhibition of mitosis would appear to contradict the concept of sensitive volumes, but more careful consideration shows this not to be the case at all. If the initiation of the mitotic process requires the participation of a certain minimum number of structural units, e.g., 7 out of 14 centromeres or genes in the diploid, the same number 7 may be required in the polyploids and not one-half the total number. Analogous situations are common in the phenotypic expression of allelic genes where the presence of a single dominant factor suffices to bring about the same phenotypic expression as two or more such alleles. The radiation will reduce the number of active mitotic centers in proportion to the number present. Thus if one-half are inactivated by any given dose, this will leave 7 active in the diploid, 14 in the tetraploid, etc. The dose must therefore be increased in the tetraploid in order to reduce the number of active centers to the point where mitosis is inhibited. It will be clear therefore that this case cannot be analyzed in terms of sensitive volumes, where the basic assumption for the necessary calculations is that each functional unit contribute to the process under consideration.

It is of interest to note that the diameter of the sensitive volume 2.2×10^{-6} cm. for *T. monococcum*, which has the basic chromosome number, is close to that calculated by Crowther[6] for chick fibroblasts, i.e., somewhat less than 4×10^{-6} cm.

Further investigation of this problem is in progress.

Summary.—1. Inhibition of mitosis by x-rays is inversely proportional to the number of chromosomes in the species studied.

2. Inhibition of mitosis is neither directly nor inversely proportional to chromosome length.

3. The evidence indicates that inhibition of mitosis by x-rays occurs through action upon the centromeres.

4. The significance of these findings in connection with sensitive volume calculations is discussed.

The writers wish to express their thanks to Dr. F. N. Briggs for seed

of the *Triticum* species and to Dr. G. L. Stebbins for seed of the *Bromus* species.

[*] Supported by a grant from the Wolf Fund administered by the Committee on Medical Research of the University of California.

[1] Strangeways, T. S. P., and Oakley, H. E. H., *Proc. Roy. Soc. London*, **B95,** 373–381 (1923).

[2] Strangeways, T. S. P., and Hopwood, F. L., *Ibid.*, **B100,** 283–293 (1926).

[3] Henshaw, P. S., and Francis, D. S., *Biol. Bull.*, **70,** 28–35 (1936).

[4] Henshaw, P. S., *Am. Jour. Cancer*, **33,** 258–264 (1938).

[5] Crowther, J. A., *Proc. Roy. Soc. London*, **B96,** 207–211 (1924).

[6] Fröier, K., Gelin, O., and Gustafsson, Å., *Bot. Notiser*, Lund (1941).

[7] Marshak, A., *Proc. Nat. Acad. Sci.*, **23,** 362–369 (1937).

[8] Yule, G. U., and Kendall, M. G., Charles Griffin Co., London (1937).

[9] Marshak, A., *Proc. Nat. Acad. Sci.*, **25,** 502–510 (1939).

[10] Marshak, A., *Proc. Soc. Exp. Biol. & Med.*, **41,** 176–180 (1939).

[11] Exner, F. M., and Luria, S. E., *Science*, **94,** 394–395 (1941).

[12] Lacassagne, A., and Wollman, E., *Compt. Rend. Soc. de Biol.*, **131,** 857–859 (1939).

EXPERIMENTS ON SEXUAL ISOLATION IN DROSOPHILA I. GEOGRAPHIC STRAINS OF DROSOPHILA WILLISTONI

BY TH. DOBZHANSKY AND ERNST MAYR

COLUMBIA UNIVERSITY AND AMERICAN MUSEUM OF NATURAL HISTORY, NEW YORK

Communicated July 11, 1944

If mature and sexually active females and males of two or more animal species inhabit the same portion of space they do not mate at random. For various reasons females of each species mate with and are inseminated by males of their own species more often than by males of other species. Causes which make mating and insemination selective rather than random may be grouped under the categories of sexual and mechanical isolation. Sexual (psychological) isolation includes various differences in behavior, in smells, and in adornments which make females and males of different species less likely to find each other, and if found less attractive and therefore less willing to accept the courtship and invitation to mating. Mechanical isolation refers to differences in the structure or function of the genitalia and other body parts which play a rôle in the act of copulation and which handicap normal intromission and sperm delivery. The distinction between sexual and mechanical isolations is to a certain extent arbitrary, and it is probable that with more study situations will be found which will be intermediate or will require a different classification.

Aside from sexual and mechanical isolation, there are other causes intrinsic to the organisms concerned which handicap or prevent the gene exchange between populations of different species. We have proposed to designate all such causes reproductive isolating mechanisms, and have reviewed them elsewhere.[1-4] Gene exchange between pairs of species is prevented usually not by a single but by several isolating mechanisms; each mechanism by itself is frequently inadequate to maintain the species separation but they do so when combined. Furthermore, the separation of different pairs of species, even within the same genus, is often caused by different sets of isolating mechanisms. It must be emphasized in this connection that a complete isolation of species is not necessarily accompanied by hybrid sterility. Unquestionably good species, including such that are classified in different subgenera or genera, are sometimes not separated by sterility barriers. The Mallard (*Anas platyrhynchos*) and the Pintail ("*Dafila*" *acuta*), the two most common ducks of the temperate zone of the northern hemisphere, coexist without mixing over many millions of square miles even though the F_1, F_2, and backcross hybrids are completely fertile (Phillips[5]). Only a few natural hybrids between these species have been found (less than one in 50,000 killed). Gordon[6] found a high fertility of hybrids but absence of interbreeding in nature between sympatric species of the fish genera *Platypoecilus* and *Xiphophorus*. We believe that sexual isolation is the most important of isolating factors in many groups of animals (mammals, birds, most insects). This is not because sexual isolation prevents the gene exchange between species any more securely than do other isolating mechanisms, such as, for example, inviability of hybrids and hybrid sterility. The advantage of sexual isolation is on a different plane: it permits two or more species to occupy the same general territory without a reproductive wastage of producing poorly viable or sterile individuals. This consideration suggests that natural selection might favor the establishment of sexual isolation as early as possible during the process of speciation.

One of the first questions that arises when we consider sexual isolation or any other isolating mechanism is whether it occurs only between species or, in at least rudimentary form, between geographic races as well. The answer to this question may clear up the problem of the origin of isolating mechanisms. If the theory is valid that species emerge gradually from geographic races then indications of incipient isolation should occasionally be encountered between geographic races, exactly as all other differences between species are foreshadowed in this manner. Dobzhansky and Koller[7] found that the races of *Drosophila miranda* which inhabit the state of Washington and the Sierra Nevada of California respectively show a weak preference for homogamic mating, and the same has been found by Stalker[8] to be true with respect to strains of *D. americana* coming from different localities in the state of Ohio. Matings among members of the same strain, race, or species shall be referred to as *homogamic* and those between different strains, races, or species as heterogamic. The method used by these authors for investigating sexual isolation is the so-called multiple choice technique: one kind of males is placed in a container with two kinds of suitably marked females, and it is tested whether and to what extent the insemination is at random. This is determined by dissection of the females after a certain amount of time has elapsed and by examination of their seminal receptacles to see whether or not they contain sperm. Provided that only a part of the females have been inseminated, it is easy to tell whether the insemination of the two kinds of females takes place selectively or at random. The method has the drawback that it gives a joint estimate of the magnitude of sexual and of mechanical isolations. Theoretically, the same frequency of insemination may occur because males of the strain A do not attempt to copulate with females of B, or because the copulation is not completed on account of an incongruence of the genitalia. Although mechanical difficulties for copulation are probably not encountered in crossing races or closely related species of Drosophila, the method gives no information about which particular link in the chain of processes leading from courtship through copulation to insemination may be responsible for the selective insemination obtained. Finally, although it is easy to determine whether or not a female has some sperm in its seminal receptacle or its spermathecae, it is difficult to estimate the quantity of the sperm. It is possible that in some copulations the amount of the sperm delivered is less than in others. Patterson, Stone, and Griffen in *D. virilis* and *D. americana*, Stone in *D. hydei*, Griffen in species of the *D. melanica*, and Mainland in those of the *D. funebris* group recorded the proportions of pair matings which produce offspring under uniform laboratory conditions.[9] They found that in some instances this proportion is greater when the female and the male mated are derived from the same than when they come from different geographic strains. It is probable that this is due in part to incipient sexual isolation between geographic races of the same species, although the production of offspring evidently depends on many more variables than insemination does.

D. willistoni Sturtevant is a species common in tropical America; its known distribution extends from the southern tip of Florida to the southern part of the state of São Paulo, i.e., from just north of the Tropic of Cancer to just south of the Tropic of Capricorn. We have living strains which fit the description of *D. willistoni* from Praia Grande and Bertioga (state of São Paulo), from Rio de Janeiro, Belem (state of Pará), Teffé (state of Amazons) in Brazil, and from Quiriguá in Guatemala (Fig. 1). With the possible exception of Quiriguá, each of these strains is derived from several females caught in the respective localities; the Quiriguá strain has been obtained through the courtesy of Professor A. H. Sturtevant. The Teffé strain fails to cross, except very rarely, with any of the other. This situation will be dealt with in another communication, and here we are concerned with the other five strains.

The experiments were arranged as follows. Ten virgin females of a given strain are placed in a $3^3/_4 \times 1$-inch vial, in which a small amount of the cornmeal-molasses-agar medium was poured and allowed to solidify in a slanting position; ten virgins[10] of another strain and ten freshly hatched males of one of these strains are added. Since the flies from the different strains are not distinguishable in appearance, the right wings are clipped off in one and the left wings in the other class of females. The vials are put in an incubator at $24^1/_2$°C. for 40–55 hours, whereupon all the females are dissected and their tubular receptacles and spermathecae examined under a microscope for sperm. Fertilized females have spermatozoa as a rule both in the receptacle and in the spermathecae, only a few individuals having been found showing sperm in the receptacle only. For statistical purposes it is best to have approximately one-half of all females inseminated at the time of the dissection; this time was adjusted accordingly, but some dissections have also been made of females which were exposed to the males for 24 and 72 hours. The data obtained in 24-

and 72-hour dissections serve as a check on the possibility that there is a significant excess of homogamic or heterogamic matings among the first or the last flies fertilized in a culture. This possibility is not realized.

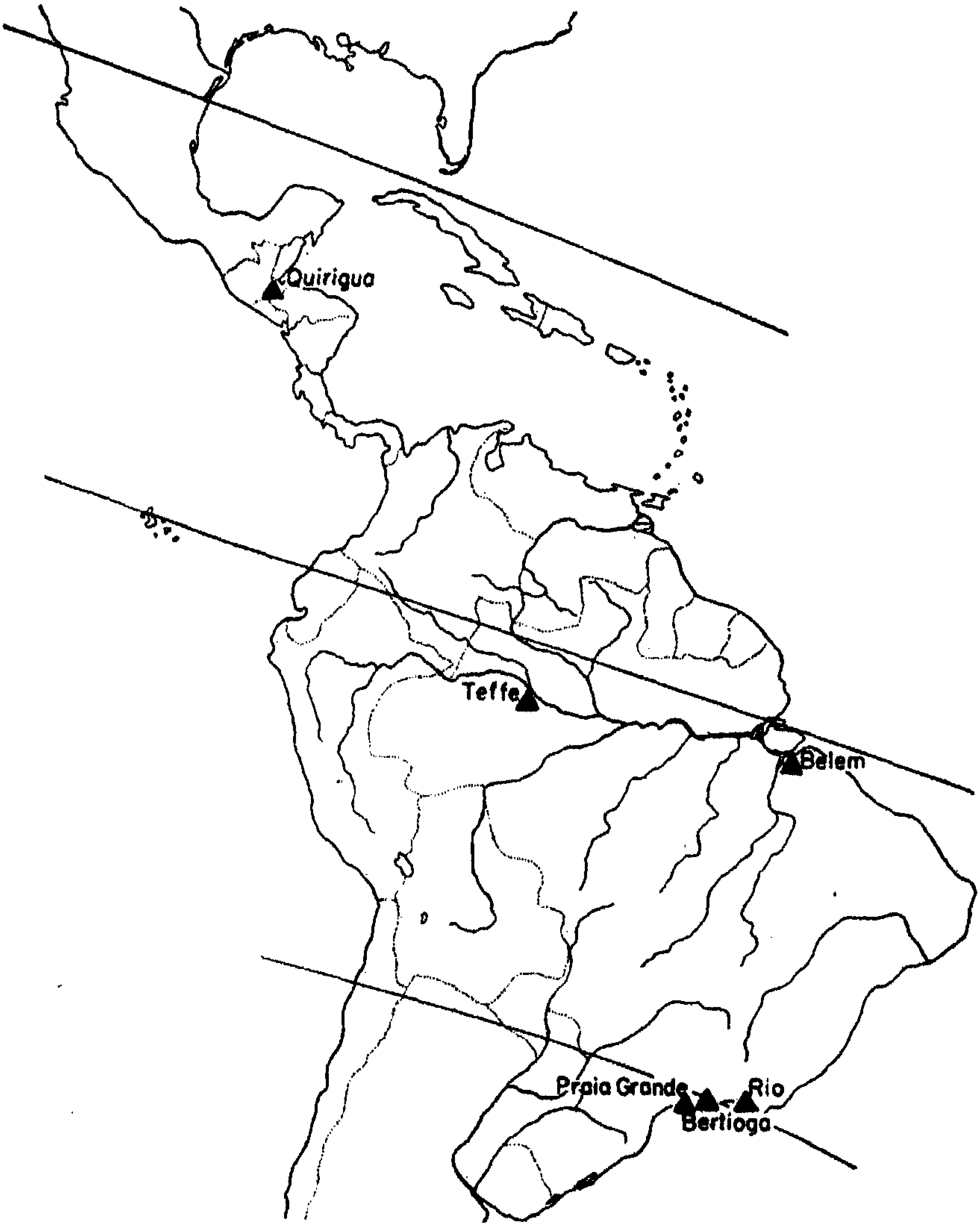


FIGURE 1
A map showing the localities in which the strains of *D. willistoni* mentioned in the text have been collected.

. A summary of the data is presented in table 1 (except that the 24- and 72-hour dissections and the occasional other cultures in which too few or too many females were fertilized are not included). In this table the percentages of homogamically and heterogamically fertilized females are given for every cross, together with the χ^2's which measure the statistical significance of the differences observed between the frequencies of the homogamic and heterogamic matings, and with the "isolation indices." This last value has been proposed by Stalker[8] to measure the degree of randomness of the matings; the isolation index equals the percentage of homogamic minus the percentage of heterogamic inseminations, divided by the sum of these percentages. If the isolation between the forms tested is complete and only homogamic matings occur the index is +1.0; if the matings occur at random the index is zero, and it is negative if there is a preference for matings between representatives of different strains.

TABLE 1

NUMBER OF FEMALES DISSECTED (*n*) AND PER CENT CARRYING SPERM (%) IN VARIOUS CROSSES OF *D. willistoni*

FEMALES	MALES	HOMOGAMIC *n*	HOMOGAMIC %	HETEROGAMIC *n*	HETEROGAMIC %	χ^2	ISOLATION INDEX
Praia Grande, Bertioga	Bertioga	57	54.4	57	57.9	0.10	−0.03
Praia Grande, Bertioga	Praia Grande	74	66.2	73	65.8	0.003	0.00
Praia Grande, Rio	Rio	69	60.9	71	69.0	1.02	−0.06
Praia Grande, Rio	Praia Grande	59	66.1	64	56.2	1.29	0.08
Bertioga, Belem	Belem	73	60.3	78	56.4	0.16	0.03
Bertioga, Belem	Bertioga	73	71.2	78	52.5	5.49	0.15
Rio, Belem	Belem	69	65.2	71	62.0	0.16	0.03
Rio, Belem	Rio	64	50.0	60	58.3	0.88	−0.08
Brazilian	Brazilian	538	62.1	552	59.8	0.61	0.02
Quiriguá, Rio	Rio	57	57.9	60	40.0	3.70	0.18
Quiriguá, Rio	Quiriguá	59	39.0	63	41.3	0.06	−0.03
Quiriguá, Belem	Belem	68	48.5	74	25.7	7.97	0.31
Quiriguá, Belem	Quiriguá	64	45.3	60	48.3	0.13	−0.03
Quiriguá, Bertioga	Bertioga	77	66.2	69	20.3	31.02	0.53
Quiriguá, Bertioga	Quiriguá	67	26.9	65	35.4	1.11	−0.14
Quiriguá, Praia Grande	Praia Grande	65	75.4	66	56.1	5.48	0.15
Quiriguá, Praia Grande	Quiriguá	90	24.4	92	62.0	25.87	−0.44

Table 1 shows that no significant deviations from randomness of mating are observed between any of the Brazilian strains (Praia Grande, Bertioga, Rio, Belem). The highest χ^2 value for such matings is 5.49, which, for 1 degree of freedom, indicates the probability of about 0.02 of the deviation being due to chance. Summing up the results of the Brazilian crosses, we have 62% of homogamic and nearly 60% of heterogamic inseminations among 1090 females dissected, the difference being statistically non-significant. A different situation obtains when the Guatemalan strain (Quiriguá) is crossed to the Brazilian ones. The isolation index is positive whenever the males used are Brazilian, and negative when Quiriguá males are used. This suggests that when a mixture of Brazilian and Guatemalan females is offered either to Brazilian or to Guatemalan males, Brazilian females are inseminated in preference to the Guatemalan ones. To be sure, the observed deviations from random insemination are mostly not significant statistically (the χ^2 values corresponding to the 0.05, 0.02, and 0.01 probability levels are 3.84, 5.41 and 6.64, respectively). An unquestionably significant positive isolation index is observed only when Bertioga males are offered a choice of Bertioga and Quiriguá females, and a negative isolation index only in the cross of Quiriguá and Praia Grande females to Quiriguá males. Combining the data for the crosses in which Brazilian males were offered a choice of Brazilian and Guatemalan females, we find the frequencies of the homo- and heterogamic matings to be 62.2% and 34.9%, respectively, with a χ^2 of 39.80, which is very significant. A similar summation of the data for the crosses of Guatemalan males to mixtures of Guatemalan and Brazilian females gives 32.9% for homogamic and 48.2% for heterogamic matings; χ^2 is 13.68 corresponding to about the 0.001 probability level.

TABLE 2

Number of Females Dissected (n) and Per Cent Carrying Sperm (%) in Crosses made in Quart-Sized Bottles

FEMALES	MALES	HOMOGAMIC		HETEROGAMIC			ISOLATING INDEX
		n	%	n	%	χ^2	
Bertioga, Belem	Bertioga	107	73.8	107	69.2	0.57	0.03
Quiriguá, Rio	Rio	192	54.2	198	39.4	4.55	0.16
Quiriguá, Praia Grande	Quiriguá	101	49.5	105	76.2	15.29	−0.21

The degree of preference for homo- or heterogamic mating is probably subject to environmental modification. An unpublished experiment of Professor W. P. Spencer suggested that the amount of space available to the flies may influence the degree of mating preference (personal communication). To test the results of the main body of experiments, in which 30 flies were confined in $3^{3}/_{4}$- × 1-inch vials, some crosses were repeated using 40 females of each of two strains and 40 males of one of them in quart milk bottles. Some culture medium was poured on the bottom of the bottles, and crumpled paper toweling was placed so that the entire space in the bottle was easily accessible to the flies. The results are summarized in table 2. They fully confirm the conclusions reached previously. No significant deviation from randomness of mating is observed when Bertioga males are given a choice of Bertioga and Belem females (Brazilian × Brazilian mating). Brazilian (Rio) males inseminate more Brazilian than Guatemalan females. Guatemalan males likewise inseminate more Brazilian than Guatemalan females.

The writers take the pleasure of acknowledging their obligations to Miss Irene Markreich for her assistance in conducting the experiments.

Summary.—With the exception of the Teffé strain which will be discussed in another publication, no trace of sexual isolation is observed between strains of *D. willistoni* from different parts of Brazil and from Guatemala. When Brazilian males are offered a choice of females of their own and of other Brazilian strains, inseminations occur at random. If a mixture of Brazilian and Guatemalan females are confined either with Brazilian or with Guatemalan males, a greater proportion of Brazilian than of Guatemalan females are inseminated. The mating of Brazilian and Guatemalan flies is, therefore, selective rather than random; however, the particular type of selectivity here observed does not constitute a barrier to gene exchange.

[1] Dobzhansky, Th., *Amer. Natur.*, **71**, 404–420 (1937).

[2] Dobzhansky, Th., *Genetics and the Origin of Species*, New York, 1937; 1941.

[3] Mayr., E., *Systematics and the Origin of Species*, New York, 1942.

[4] The expression "physiological isolating mechanisms" used in some papers of Dobzhansky was intended to distinguish between isolating mechanisms based on inherent properties (physiology) of the organisms concerned and geographical isolation. Since this expression has led to some misapprehension, the adjective "reproductive," instead of "physiological," is used to convey the same idea.

[5] Phillips, J. C., *J. Exp. Zool.*, **18**, 69–146 (1915).

[6] Gordon, M., *Trans. N. Y. Acad. Sci.*, **5** (2), 63–77 (1943).

[7] Dobzhansky, Th., and Koller, P., *Biol. Zentr.*, **58**, 589–607 (1938).

[8] Stalker, H. D., *Genetics*, **27**, 238–257 (1942).

[9] Patterson, J. T., Stone, W. S., and Griffen, A. B., *Univ. Texas Public.*, **4228**, 162–200 (1942); Stone, W. S., *Ibid.*, 16–22; Griffen, A. B., *Ibid.*, 68–73; Mainland, G. B., *Ibid.*, 74–112.

[10] Copulation in *D. willistoni* begins as early as 8 hours after the hatching from pupae; to collect virgins bottles were emptied at 4-hour intervals, but it was found expedient to use also females hatched overnight since the proportion fertilized is small.

THE COMPRESSIBILITY OF MEDIA UNDER EXTREME PRESSURES

By F. D. Murnaghan

Department of Mathematics, The Johns Hopkins University

Communicated July 11, 1944

It is well known that Hooke's Law, which postulates a linear relation between stress and strain, has a very limited range of applicability even when the applied stress is a uniform pressure. We have in previous papers[1,2] furnished a formula which is valid over a much greater range than Hooke's Law; this formula agreed well with experimental results up to pressures as high as 50,000 atmospheres (the highest for which measurements were then available). Since then Bridgman[3] has published measurements up to 100,000 atmospheres and this, combined with theoretical considerations of a fundamental character, has caused us to reconsider the whole question. Hooke's Law and our formula may both be derived from the principle of energy conservation; the difference in the results is due to the fact that for infinitesimal deformations (to which the validity of Hooke's Law is limited) it is correct to say that stress is the gradient, with respect to strain, of the energy per unit volume. For non-infinitesimal deformations this simple relation must be replaced by a more complicated one. The feature common to both theories is the *assumption* that the energy density is a function of the strain; it being agreed that the strain is measured from the position of zero stress. This assumption, which has apparently never been seriously questioned, has the quality of an "action at a distance" theory; we assume that the energy of deformation is furnished by a knowledge of the relationship of the actual position of the medium to a *remote* position in which the medium was unstressed. We propose here to discard this action at a distance theory and to replace it by a *differential* theory in which we are concerned merely with the variation of the energy as we pass from *any* position of the medium to an infinitesimally near position. Thus we must be prepared to confront the situation where the *initial* position of the medium is one in which the medium is under stress.

This being understood, a fundamental question arises: are the elastic *constants* of the medium really constant? The elastic constants are simply coefficients which occur in the statement of Hooke's Law and we mean by the word *constant* that they do not depend on the *increment* of the stress. But do they depend on the initial value of the stress? All experimental evidence points to the fact that they do, and many determinations have been made of the variation of such things as compressibility with pressure. It is fair to say that the reason that Hooke's Law is so limited in its range of applicability is that it neglects the dependence of the elastic constants on the initial stress. This dependence of the elastic constants on the initial stress cannot be overemphasized. For example, when we say that an elastic medium is isotropic, we mean that certain relations hold amongst the elastic constants (so that *two* are sufficient to furnish the statement of Hooke's Law); it is clear, then, that a medium which is elastically isotropic under zero stress may fail to be elastically isotropic when stress is applied. We should certainly expect this lack of isotropy to appear when the applied stress is (as in the case of the Young's modulus experiment) not merely a uniform pressure or traction. We feel certain that it is the reliance of the classical theory of elasticity upon the hope that a stretched wire is elastically isotropic, if the unstretched wire is, that is responsible for its complete failure to predict the yield point phenomenon.

The case of uniform pressure is particularly simple since the stress tensor is scalar. The principle of conservation of mass, combined with Hooke's law for *infinitesimal* variations of stress, suffices provided we are willing to face the situation that the compressibility depends on the pressure. The principle of mass conservation yields $\frac{1}{v}\,\delta v = 3\,\frac{\partial}{\partial x}\,\delta x$ and Hooke's Law yields $p = -(3\lambda + 2\mu)\,\frac{\partial}{\partial x}\,\delta x$ and so

$$\frac{1}{v}\frac{dv}{dp} = -\frac{1}{\lambda + {}^2/_3\mu}.$$

Let us assume, as a first approximation, that $\lambda + {}^2/_3\mu$ is a linear function of p:

$$-v\frac{dp}{dv} = c(1 + kp)$$

so that

$$c = -\left(v\frac{dp}{dv}\right)_{p=0}; \qquad ck = -\frac{d}{dp}\left(v\frac{dp}{dv}\right)_{p=0}.$$

We find $\frac{v_0}{v} = (1 + kp)^{1/ck}$ or, on writing $\Delta v = (v_0 - v)$,

$$\frac{\Delta v}{v_0} = 1 - (1 + kp)^{-1/ck}.$$

This simple formula accounts reasonably well for the experimental results even for such a range of pressure as from 0 to 100,000 atmospheres. The data for Li given by Bridgman (loc. cit.) are approximated by setting $k = 0.153 \times 10^{-4}$, $ck = 2$ so that the relation connecting p and v has the simple form

$$0.153 \times 10^{-4}p = \left(\frac{v_0}{v}\right)^2 - 1; \qquad \frac{\Delta v}{v_0} = 1 - (1 + 0.153\cdot10^{-4}p)^{-1/2}.$$

Bridgman tabulates $\frac{\Delta v}{v_0}$ at intervals of 10,000 atmospheres. The following table furnishes the comparison between the values obtained empirically and those calculated from the formula just given

$p =$	10^4;	2.10^4;	3.10^4;	4.10^4;	5.10^4;	6.10^4;	7.10^4;	8.10^4;	9.10^4;	10^5
$\frac{\Delta v}{v_0}$ (obs.)	0.074;	0.125;	0.164;	0.201;	0.237;	0.272;	0.305;	0.336;	0.366;	0.394
$\frac{\Delta v}{v_0}$ (calc.)	0.069;	0.125;	0.172;	0.212;	0.247;	0.278;	0.305;	0.330;	0.353;	0.371

The formula predicts that when $p = 2.10^5$ $\frac{\Delta v}{v_0}$ will be approximately 0.514; when $p = 10^6$ $\frac{\Delta v}{v_0}$ will be approximately 0.752. To compress Li to a density ten times its original density, a pressure of around 6.5×10^6 atmospheres is necessary.

In conclusion we add the remark that when the pressure is small enough so that its influence upon the elastic constants is negligible the appropriate formula is

$$\frac{v_0}{v} = e^{p/c}; \qquad \frac{\Delta v}{v_0} = 1 - e^{-p/c}.$$

ADDED IN PROOF. It is worthy of attention that the formula given here is, when the dimensionless constant ck is assigned the value 2, precisely that which leads to Laplace's well known law of density distribution throughout the interior of the earth. This law fits the observed facts remarkably well but it has been regarded by many authorities as merely empirical largely because its connection with the law of compressibility has not been clearly presented.

---

[1] "Finite Deformations of an Elastic Solid," *Amer. Jour. Math.*, **59**, 235–260 (1937).

[2] "The Compressibility of Solids," *von Kármán Anniversary Volume*, 1941, pp. 121–136.

[3] "Compressions and Polymorphic Transitions of Seventeen Elements to 10^5 Kg./Cm.^{2}," *Phys. Rev.*, **60**, 351–354 (1941).

NOTE ON THE DISCOVERY OF RED STARS

BY GUILLERMO HARO

HARVARD COLLEGE OBSERVATORY

Communicated July 26, 1944

During an investigation of colors, magnitudes, and spectral types in the Hercules-Vulpecula region, it was noted that a star at the approximate position $19^h 10^m.6$, $+21° 37'$ (1855), $\lambda = 23°$, $\beta = +3°$ had a very large color index. A preliminary examination of blue and red plates indicated that the blue-red color index was between 6.5 and 7.0 magnitudes.

The blue-red and blue-yellow (international) color indices of this star were determined from polar comparison plates, the blue magnitude from Cramer Hi-Speed plates, the red magnitude from Eastman 103a-E plates with a ciné-red filter (effective wave-length near 6300 A), the yellow magnitude from 103a-G plates with a yellow filter. The 8-inch Ross-Lundin (IR) refractor was used for the red and the yellow, the 16-inch Metcalf (MC) refractor for the blue magnitudes. All plates were taken between May 26 and June 5, 1944. The results were as follows:

Mag$_{pg}$	$14\overset{m}{.}75$
Mag$_{pr}$	$8\overset{m}{.}10$
Mag$_{pv}$	$9\overset{m}{.}75$
C.I. (blue-red)	$6\overset{m}{.}65 \pm 0.10$
C.I. (blue-yellow)	$5\overset{m}{.}00 \pm 0.10$

This star has been observed before by Lee and his associates; it is No. 129 in the list of Dearborn Annals, Vol. IV, Part 16, 1940, where its spectral type is given as *N*.

The star is not unlike the very red star discovered in 1935 by F. K. Edmondson and A. M. Rogers.[1] Both Edmondson's star, BN Monocerotis, and the present object are variable.[2] An examination of plates in the Harvard Collection gives the following provisional results:

DATE	MAG$_{pg}$
Aug. 31, 1897	13.55
Sept. 11, 1917	15.10
July 5, 1926	13.80
Oct. 28, 1927	14.30
Apr. 18, 1928	14.35
Oct. 9, 1928	14.10
June 3, 1929	13.70
July 15, 1929	13.60
July 18, 1930	14.00
May 11, 1937	13.90
July 13, 1937	13.70
May 28, 1938	13.60
May 30, 1938	13.60
Aug. 30, 1940	14.60
May 26, 1944	14.80
May 28, 1944	14.85
May 29, 1944	14.75

The finding of this very red star led the author to attempt a search for other very red stars in the same region of the sky. The method employed in this search resembles that used previously by several astronomers, among others, G. A. Tikhov,[3] V. M. Slipher[4] and G. Z. Dimitroff.[5]

The instrument used in the survey was again the 8-inch Ross-Lundin (IR) refractor. The region was photographed on an Eastman 103a-E plate, exposure time five minutes at normal blue focus and without filter. Since the IR camera has a very steep focal curve, ($^1/_5$ inch distance between blue and red focus) and the 103a-E plates are sensitive to the blue as well as to the red region of the spectrum, the red stars appear as a black dot (blue) surrounded by a strong red halo, when focused for the blue. The halo is almost totally absent for the very blue stars, but quite conspicuous for the normal late-type stars. The above-mentioned star is so red that the blue central dot is absent on a five-minute exposure; there appears on the plate only a halo of fairly uniform brightness with a central hole.

The region was photographed a second time with the IR camera, blue focus, 103a-E plates, no filter, and with an exposure time of thirty minutes. This plate revealed the presence of three additional very red stars, marked by a red halo without the central blue image. The approximate positions, magnitudes and color indices of these stars are:

	STAR 2	STAR 3	STAR 4
R.A. (1855)	$19^h\ 9^m.2$	$19^h\ 8^m$	$19^h\ 9^m$
Dec. (1855)	+24° 58′	+18° 20′	+17° 10′
λ	26°	18°	17°
β	+5°	+3°.5	+3°
Mag_{pg}	$16^m.40$	$15^m.60$	$15^m.70$
Mag_{pr}	9.90	10.45	10.35
Mag_{pi}	11.50		
C.I. (blue-red)	$6.50 \pm 0^m.10$	$5.15 \pm 0^m.10$	$5.35 \pm 0^m.10$
C.I. (blue-yellow)	4.90 ± 0.10		

The spectral type of the second star was determined from two plates taken by Miss Florence Anderson with the 12-inch Metcalf (MA) refractor, exposure time two hours. The spectrum is faint, and resembles that of an *N* star.

Following the search of the first field, discovery plates were made for two additional centers, one eight degrees north, the other eight degrees south of the first center, with thirty minutes' exposure time. An examination of these two plates reveals the presence of four additional conspicuously red stars, with a red halo, but no central blue image. It seems likely that the blue-red color index of each of these stars is well in excess of four magnitudes. The approximate 1855 positions from BD-charts are:

STAR	R.A. (1855)	DEC. (1855)	λ	β
5	$18^h\ 58^m.3$	+13° 30′	13°.5	+2°.5
6	19 11	+ 8 35	9	−1.5
7	19 10	+27 36	28	+6.5
8	19 16.3	+31 2	31	+7

The total area covered by the present survey is approximately 250 square degrees; it is complete to the eleventh photo-red magnitude. The survey yields eight stars with color indices (blue-red) of more than four magnitudes. It should be noted that all stars are within eight degrees of the galactic circle. A study of color indices of stars of known spectral type shows the average absorption for the region to be of the order of one magnitude per kiloparsec. None of the stars is found in a region of heavy obscuration included in the survey.

[1] *Pub. A.S.P.*, **47**, 191 (1935).

[2] *Ap. Jour.*, **100**, 1 (1944).

[3] *A.N.*, **218**, 145 (1922) and *Poulkovo Publ.*, Ser. 2, **50** (1937).

[4] *Pub. A.A.S.*, **7**, 15 (1931).

[5] *The Telescope*, **6**, 62 (1939).

---

*ANTIBIOTIC ACTIVITY OF LICHENS**

BY PAUL R. BURKHOLDER, ALEXANDER W. EVANS, ILDA MCVEIGH AND HELEN K. THORNTON

OSBORN BOTANICAL LABORATORY, YALE UNIVERSITY

Communicated July 20, 1944

The production by an organism of organic substances inimical to the life processes of other organisms and inhibiting their growth is known as antibiosis. The phenomenon of antibiosis has attracted wide-spread attention in recent years chiefly because of the remarkable chemotherapeutic effects of penicillin, a product of the mold *Penicillium*. The anabolic products of many kinds of bacteria,[1] molds,[2] and algae,[3] and substances elaborated by certain species of flowering plants[4] can be readily demonstrated to have antibiotic properties when small amounts of the active materials are tested on microörganisms with suitable assay procedures. In view of the reported antibacterial activity of the green alga *Chlorella* and the many antagonistic substances now known to be produced by numerous kinds of fungi, the lichens seemed to offer favorable material for antibiotic investigations, inasmuch as the bodies of these plants are comprised of mixtures of algae and fungi. Accordingly we have tested a considerable number of lichens for antagonistic action against some common species of bacteria, and offer herewith a preliminary report concerning some of our observations.

Methods.—The lichens were brought in fresh condition to the laboratory, placed in shallow pans, and moistened with water for the purpose of allowing them to become physiologically more active. After a few hours of exposure to sunlight, samples of the moist lichens were separated from the substrata with a razor blade and forceps. Extracts were made by grinding, with a glass mortar and pestle, 100 mg. of lichen in 1 ml. of phosphate buffer solution adjusted to pH 7.4. The final aqueous extract was approximately pH 7.0. Extractions were carried out at room temperature and the samples were stored at 1°C. for a short time until the materials could be assayed. No attempt was made to clarify the extracts in these preliminary tests. The aqueous suspensions were assayed against *Staphylococcus aureus*, *Escherichia coli*, *Bacillus subtilis* or other bacteria by means of the cylinder plate procedure.[5] As is well known to workers in this field, the method involves the use of agar plate cultures of bacteria upon which are placed small glass cups to which are added the assay liquids. After a period of incubation, zones of bacterial growth-inhibition subjacent to the cups are found to vary in diameter depending upon the potency of the test solutions which have diffused out of the cups into the agar. In these preliminary tests with aqueous suspensions the diameter of the zones of inhibition ranged up to 22 mm.

TABLE 1

ANTIBIOTIC ACTIVITY OF EXTRACTS OF LICHENS IN PHOSPHATE BUFFER TESTED AGAINST THREE BACTERIA. + = INHIBITED; – = NOT INHIBITED

SPECIES OF LICHEN	*S. aureus*	*E. coli*	*B. subtilis*
C. caespiticia	+	–	+
C. capitata	+	–	+
C. coniocraea	+	–	+
C. cristatella	+	–	+
C. glauca	+	–	+
C. Grayi	+	–	+
C. nemoxyna	+	–	+
C. squamosa	+	–	+
C. strepsilis	+	–	+
Parmelia physodes	+	–	+
Physcia sp.	+	–	+
Ramalina sp.	+	–	+
Sterocaulon paschale	+	–	+
C. incrassata	–	–	+
C. atlantica	–	–	+
C. subtenuis	–	–	+
C. caroliniana	–	–	+
C. uncialis	–	–	+
Lepra sp.?	–	–	+
Parmelia conspersa	–	–	+
Peltigera sp.	–	–	+
C. furcata	+	–	–
C. papillaria	+	–	–
Umbilicaria papulosa	+	–	–

Further experiments were performed with selected species of lichens for the purpose of determining whether the active compounds could be extracted with solvents, such as ether, chloroform and ethyl alcohol. The moist lichen materials were ground in a glass mortar, using 100 mg. of lichen to approximately 1 ml. of solvent. The suspensions were then centrifuged, and the clear solutions were pipetted into small flasks which could be attached to a vacuum distillation apparatus for removing the solvents at low temperature. The residual matter was taken up in aqueous phosphate buffer at pH 7.4, and after mixing with a glass rod the material was tested in the cup plates. Incubation of the test plates was allowed to proceed for 12 hours at 37°C.

Results.—Of the 42 species of lichens tested thus far, 27 have been found active against either *S. aureus* or *B. subtilis*. Some of these are listed in table 1. In addition to those given in the table, *Cladonia cryptochlorophaea*, *C. pyxidata* and *Dermatocarpon miniatum* were found to inhibit *S. aureus*, but were not tested with *B. subtilis*. None of the 42 lichens which were tested shows any activity against *E. coli*. Slight inhibition against *Alcaligenes fecalis* was noted for the lichen species *C. glauca* and *Parmelia conspersa*. A strain of *Proteus vulgaris* was inhibited by *Parmelia physodes* and by *Umbilicaria papulosa*. *Aerobacter aerogenes* and *Serratia marcescens* were not inhibited by the lichens which were antibiotic for *Alcaligenes* and *Proteus*. The lichens *C. atlantica*, *C. caroliniana*, *C. cristatella*, *C. Grayi*, and *Parmelia physodes*, which are active against *B. subtilis*, were tested also with *B. cereus* and *B. vulgatus*, and were found inhibitory for these species of bacteria. Species which were inactive when tested with *E. coli* and *S. aureus* are the following: *Baeomyces roseus*, *C. apodocarpa*, *C. bacillaris*, *C. clavulifera*, *C. conista*, *C. pleurota*, *C. Robbinsii*, *C. subcariosa*, *C. submitis*, *C. verticillata*, *Leptogium* sp., *P. rudecta* and *Parmeliopsis aleurites*. *Baeomyces roseus*, *C. apodocarpa* and *C. subcariosa* were inactive also with *B. subtilis*. It is possible that activity would be found if these lichens were tested against other bacteria.

The activities of different lichens which were extracted with chloroform, ethyl alcohol or ether are recorded in table 2. The equivalent weights of moist lichen material (200, 300 or 500 mg.) extracted with the solvents and finally taken up in 1.0 or 1.2 ml. of phosphate buffer solution are listed under each species in the table. Although the data are not strictly quantitative for chloroform and alcohol because of some losses incurred during transfer of the solutions, the results indicate clearly that active antibacterial substances can be extracted from lichens with convenient solvents and passed on into an aqueous test medium. That more than one antibiotic compound may exist in lichens is suggested by the fact that both *S. aureus* and *B. subtilis* are inhibited by extracts from *C. Grayi*, *Parmelia physodes* and other species, while substances obtained from some species of *Cladonia* inhibit *B. subtilis* but not *S. aureus*. Furthermore, extracts from *C. furcata*, *C. papillaria* and *Umbilicaria papulosa* inhibit *S. aureus* but are inactive against *B. subtilis*. The inhibition of some Gram-negative bacteria by selected species of lichens lends further support to the theory of multiple substances. Until more information becomes available about the nature of these antibacterial compounds, no names will be suggested for them as additions to the already confused terminology of antibiotics.

Characteristic zones of bacterial inhibition obtained with ether extracts, transferred to phosphate buffer for making the tests, are shown in figure 1. It should be mentioned here that with some lichens inhibition was obtained with both ether and aqueous extracts; in other instances antibiosis was found with only the ether or the phosphate buffer solution. As an example, an ether extract of *C. glauca* in figure 1 shows no antagonism for *B. subtilis*, but an aqueous extract made directly from the same lichen gave slight inhibitory action on the same species of bacteria. Hence we have listed *C. glauca* in table 1 as being active against *B. subtilis*.

TABLE 2

Antibiotic Activity of Lichen Extracts Tested against *B. subtilis*. Weight of Moist Lichen Is Given as Mg. per Ml. of Buffer Solution. Diameters of Bacterial Inhibition Zones Are Expressed in Mm.

SPECIES OF LICHEN	PRIMARY SOLVENT	INHIBITION ZONES
C. caroliniana	Chloroform	23
(500 mg./1.2 ml.)	Alcohol	19
	Ether	25
C. cristatella	Chloroform	26
(200 mg./1.2 ml.)	Alcohol	27
	Ether	28
C. incrassata	Chloroform	25
(200 mg./1.2 ml.)	Alcohol	25
	Ether	29
C. Grayi	Chloroform	23
(300 mg./1.2 ml.)	Alcohol	24
	Ether	25
Parmelia physodes	Chloroform	15
(200 mg./1.2 ml.)	Alcohol	15
	Ether	18
C. strepsilis (200 mg./1.0 ml.)	Ether	20
C. nemoxyna (200 mg./1.0 ml.)	Ether	20
Parmelia conspersa (200 mg./1.0 ml.)	Ether	23
Peltigera sp. (200 mg./1.0 ml.)	Ether	16
Physcia sp. (200 mg./1.0 ml.)	Ether	23
Ramalina sp. (200 mg./1.0 ml.)	Ether	21
C. uncialis (200 mg./1.0 ml.)	Ether	19

Without additional information it is impossible to say whether the antibiotic activity of lichens should be attributed to the algal or fungal member of the complex. Almost nothing is known about the anabolism of the

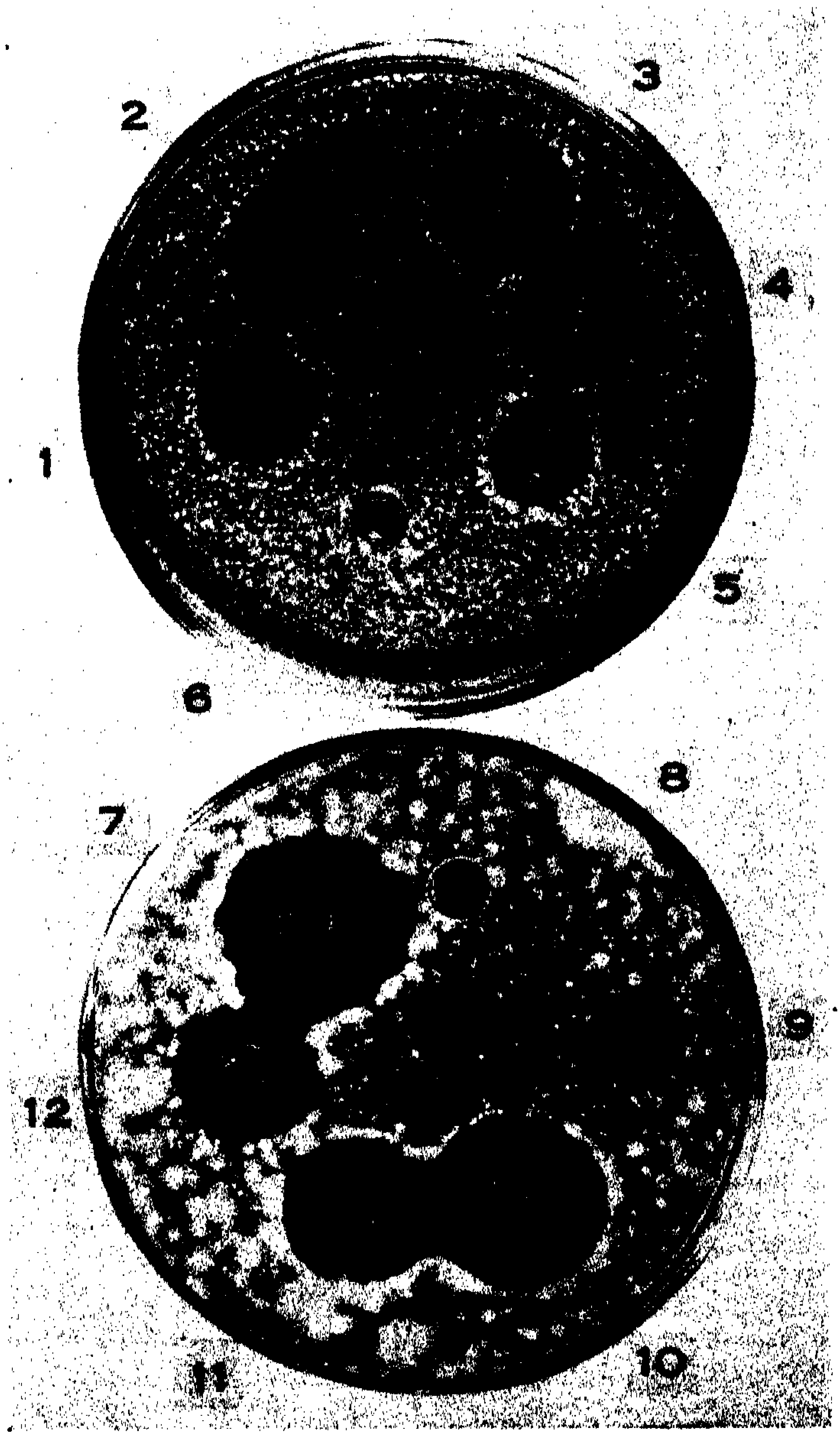


FIGURE 1

Antibacterial action of extracts of lichens. Above, lichen substances extracted with ether were tested in aqueous phosphate buffer against *B. cereus;* below, extracts tested against *B. subtilis.* Sources of the extracts: 1, *Parmelia physodes;* 2, *C. cristatella;* 3, *C. caroliniana;* 4, *C. atlantica;* 5, *C. Grayi;* 6, *C. coniocraea;* 7, *Parmelia conspersa;* 8, *C. apodocarpa;* 9, *C. glauca;* 10, *C. caroliniana;* 11, *Physcia* sp.; 12, *Ramalina* sp.

components or the roles of the various substances formed in the lichen body. Chemical antagonism would appear to constitute a possible defense mechanism operating in the struggle for existence among these primitive plants.

One question of considerable importance emerges from the data reported in this paper when considered in the light of earlier work done on the structural chemistry of the lichen acids by Zopf,[6] Asahina,[7] and others. Do the characteristic lichen acids possess antibacterial activity or are the antibiotic properties of lichens related to traces of other unidentified substances synthesized by these plants? Some of the known lichen compounds possess certain structural features in common with antibacterial substances which have been isolated from molds, but whether these are responsible for the antibiotic phenomena which we have observed can be tested only by further experiments which are in progress.

Summary.—Antibacterial activity of extracts from 42 species of lichens was determined with the cup plate procedure. Twenty-seven species were found active against *S. aureus* or *B. subtilis*, 4 species inhibited *P. vulgaris* or *Alcaligenes fecalis*, but none showed antagonism against *E. coli*.

* This work was supported in part by a research grant from Parke, Davis & Company.

[1] Dubos, R. J., and Hotchkiss, R. D., *Jour. Exptl. Med.*, **73**, 629 (1941).

[2] Waksman, S. A., *Bact. Revs.*, **5**, 231 (1941).

[3] Pratt, R., *et al.*, *Science*, **99**, 351 (1944).

[4] Osborn, E. M., *Brit. Jour. Expt. Pathol.*, **24**, 227 (1943).

[5] Foster, J. W., and Woodruff, H. B., *Jour. Bact.*, **47**, 43 (1944).

[6] Zopf, W., *Die Flechtenstoffe*, G. Fischer, Jena, 1907, 449 pp.

[7] Asahina, Y., *Flechtenstoffe. Fortschr. Chem. organischer Naturstoffe*, **2**, 27 (1939).

ACTIVITY AND NUTRITIONAL DEPRIVATION

BY GEORGE WALD AND BLANCHE JACKSON*

BIOLOGICAL LABORATORIES OF HARVARD UNIVERSITY

Communicated July 27, 1944

Animals in general exhibit increased activity when in want of food. This is in fact the indispensable sign of hunger. It has been observed in widely different organisms from mammals to protozoa;[1] in man and other animals during sleep;[2] in dogs,[3] guinea pigs,[4] and pigeons[5] after decerebration; and in rats from which virtually the entire stomach had been removed.[6] Increased activity therefore holds a primary position in the hunger reaction, persisting in higher animals which lack the equipment for either hunger sensations or the hunger contractions of the stomach.

That it has this status is not extraordinary, since in all animals feeding depends upon either moving about in the environment or, in certain aquatic forms, stirring the environment, in either case bringing the animal into contact with its food.

Sexual demands may evoke behavorial reactions similar to those associated with nutritional need. Female rats, for example, display a high peak of running when in oestrus and receptive to the male. This ceases on copulation or artificial stimulation of the cervix uteri.[7]

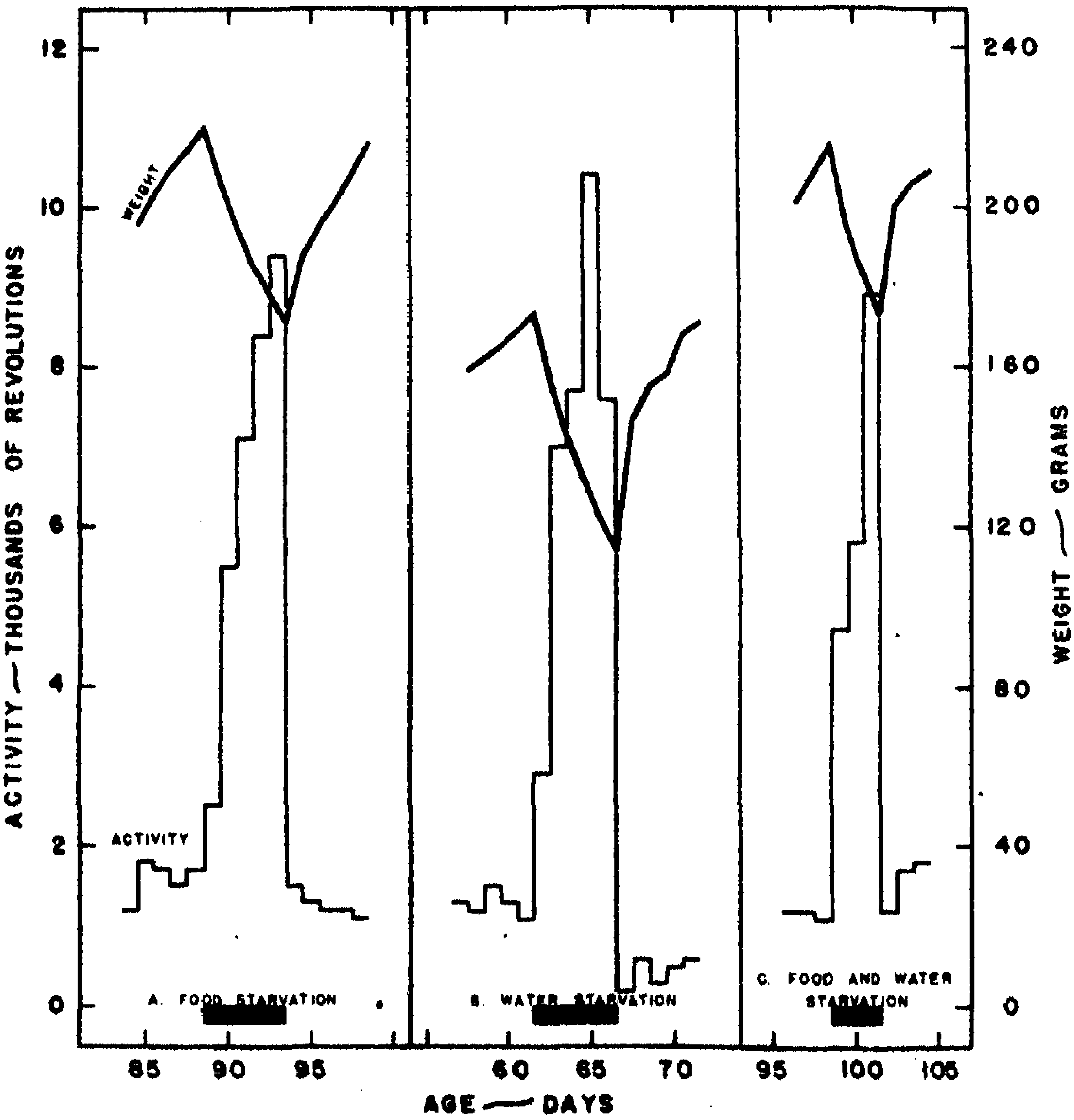


FIGURE 1

Running of the rat in food and water deprivation. Effects on the average weight and daily running of deprivation and readministration of total dry food (A, 12 animals); water (B, 3 animals); and food and water (C, 12 animals). In each figure the black area under the baseline represents the period of deprivation.

Several years ago we began experiments on the spontaneous running of white rats in nutritional deprivation. Male animals were kept in activity cages in which they had free access to a running wheel (1440 revolutions to the mile).[8] In this arrangement they normally ran an average of 0.4–1.2 miles daily.

A rat deprived of its bulk food while given water *ad libitum* greatly increases its running. On readmission of food the high running ceases and the rat begins to feed vigorously (Fig. 1 *A*). Rats allowed free access to their dry food but deprived of water react similarly (Fig. 1 *B*), as do also animals deprived of both food and water (Fig. 1 *C*).

To the degree that these responses depend upon genuine tissue needs and not simply upon mechanical factors such as the filling and emptying of the stomach, one might expect them to be elicited by single indispensable components of the diet. Water itself is such a component. Thiamin also appears to have this status.

Responses to the deprivation and readministration of thiamin are shown in figure 2.[9] They are of precisely the same form as those associated with variations in the supply of bulk food or water. Some time after our initial experiments were performed we learned that high running in thiamin deprivation had already been reported by Guerrant and Dutcher;[10] recently Bloomfield and Tainter[11] have observed similar responses in rats deprived of the entire vitamin B complex.

In rats continued on a thiamin-deficient diet maximal running persists for about 10 days. Then it falls within a further 2–8 days to far below the normal level (Fig. 3). Simultaneously the classic neuromuscular symptoms of thiamin deficiency appear: spasticity of the limbs, hunching of the back and finally acute polyneuritis.[12] Usually animals do not run below the normal level for more than a day before developing polyneuritis. The rat of figure 3, for example, in repeated trials regularly ran about 3000 revolutions on the very day before it became polyneuritic and appeared to be moribund.

From this terminal condition the animal may be rapidly restored with a single injection of thiamin. Its immediate response—beyond the cure of polyneuritis—is to eat enormously. Meanwhile it runs no more and may run even less than when polyneuritic. The weight rises rapidly for a time, then levels off. At this point the animal again begins to run, and reaches a peak of activity as eating and weight decline. Each injection of thiamin initiates such connected cycles of eating, weight-change and running, rising to maxima in that order, and returning finally to the minimal levels at which the rat is again polyneuritic (Fig. 3).

A striking interplay between thiamin and bulk food deprivation may be demonstrated as follows (Fig. 3). Food consumption falls during thiamin deficiency from normal daily values of 10–14 grams to about 2–3 grams prior to polyneuritis. If a polyneuritic animal is injected with large daily doses of thiamin and is meanwhile restricted to the low food intake it had established on the thiamin-low diet, it begins to run enormously. Thiamin, in addition to curing its polyneuritis, has restored its hunger for bulk food to a degree which the restricted allotment no longer satisfies.

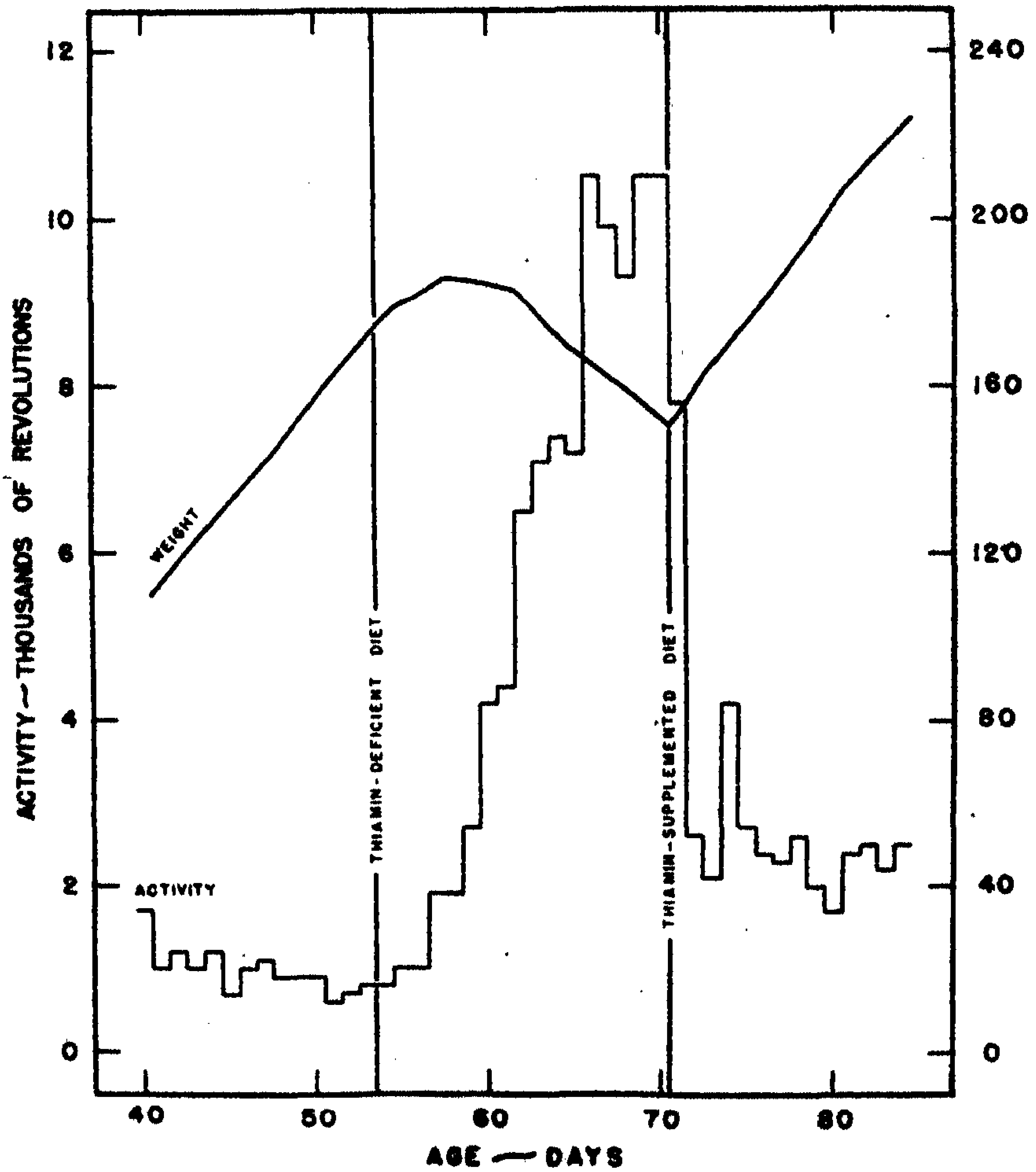


FIGURE 2

Effects of deprivation and readministration of thiamin. Averaged data from 6 animals. From 54 to 71 days of age, the thiamin supplement was omitted from an otherwise complete diet.

Specific thiamin-hunger has been replaced by the hunger for bulk food. In consequence the animal runs, and in an amount that its limited caloric intake cannot support. Having already been reduced during thiamin deficiency to an extreme state of inanition, it usually dies after a few days of great activity.

The effects of deprivation and readministration of riboflavin in rats on an otherwise complete diet are shown in figure 4. They are essentially similar to the responses already described, though not usually as prompt or marked as those associated with the thiamin nutrition.

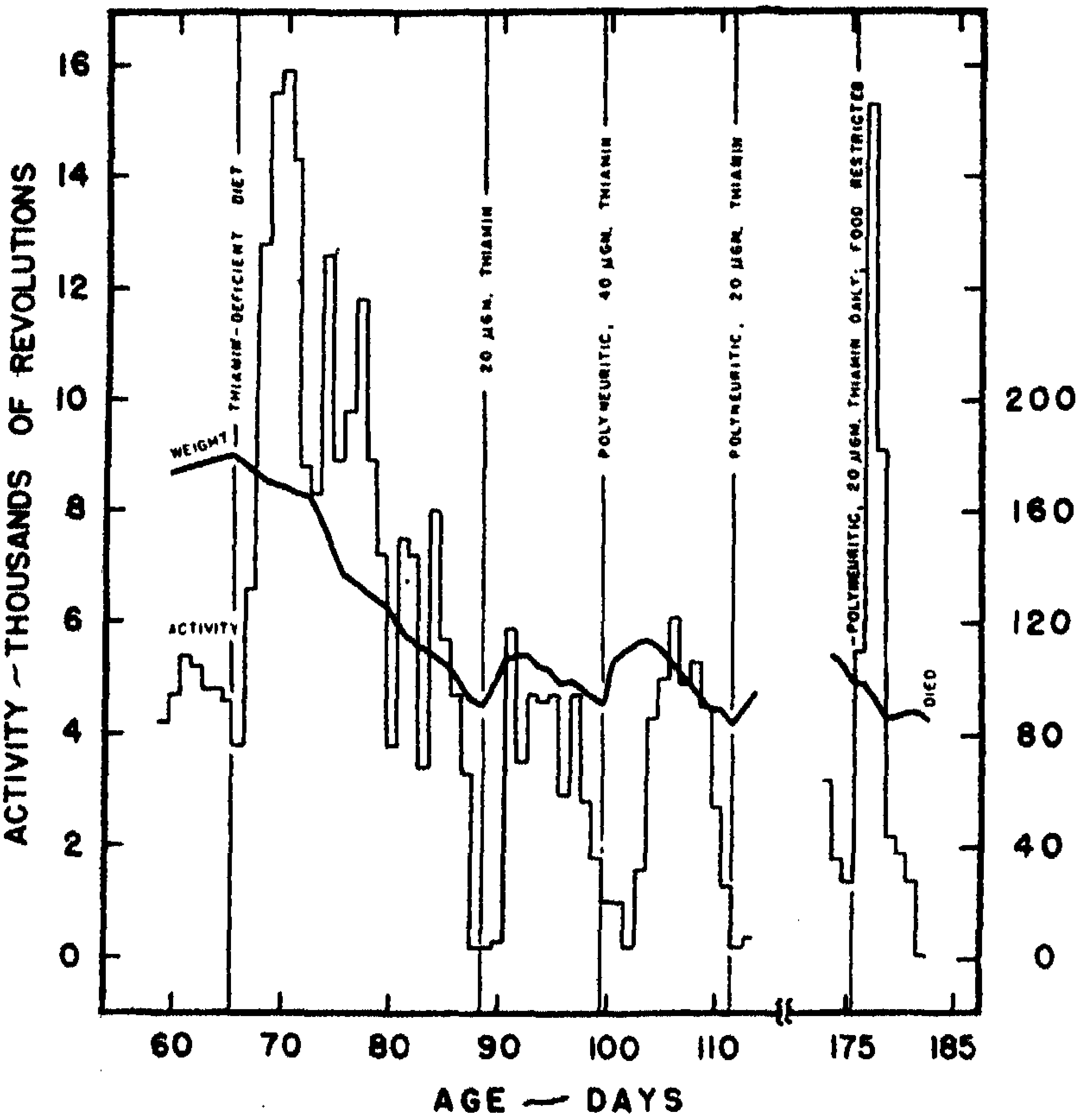


FIGURE 3

Weight and daily running of a single animal deprived of thiamin when 65 days old. On its 88th day and at intervals thereafter it was injected intraperitoneally with the amounts of thiamin indicated, being meanwhile continued on the thiamin-low ration. On the 176th day and thereafter it was injected daily with 20 μg. thiamin and was restricted meanwhile to the amounts of ration consumed while thiamin-deficient.

We have examined also the effect of depriving rats of magnesium ion (Fig. 5). In this case no high running whatever is observed. Instead the running declines after 1–2 weeks of deprivation to very low values. Magnesium is one indispensable component of the diet, therefore, the lack of which does not stimulate increased activity. This appears to be true also of deprivation in total inorganic ions,[13] or in vitamin A.[14]

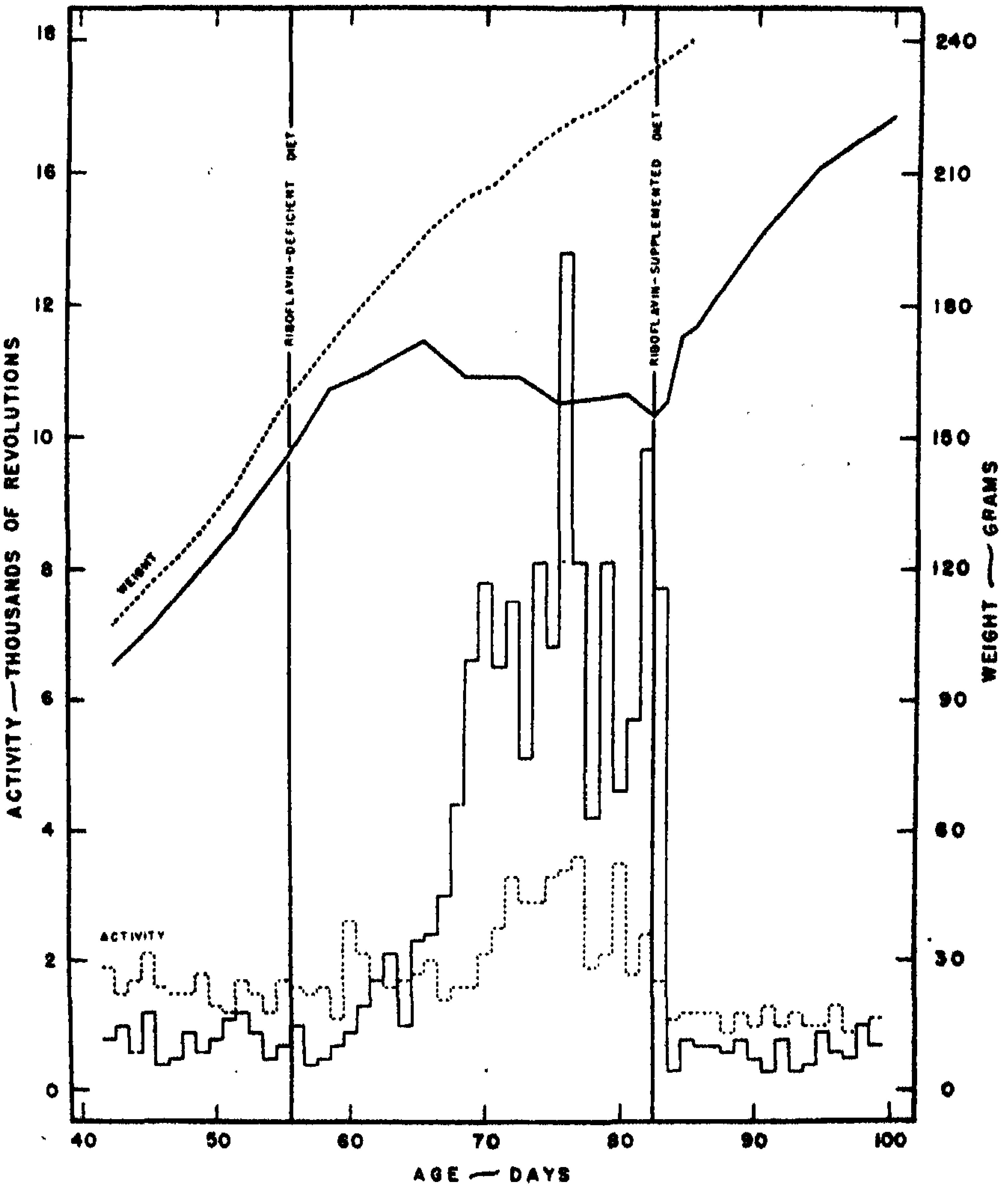


FIGURE 4

Deprivation and readministration of riboflavin. Data from a single experimental rat (solid lines) compared with averaged measurements on 5 control animals (broken lines). From age 56 to 83 days the riboflavin supplement was omitted from the diet of the experimental animal.

It may be concluded that increased running is a basic response to nutritional need, and that it may be stimulated by the lack of specific indispensable components of the diet. It is noteworthy that, except for restriction of water or of the bulk of solid food, the only types of deprivation which

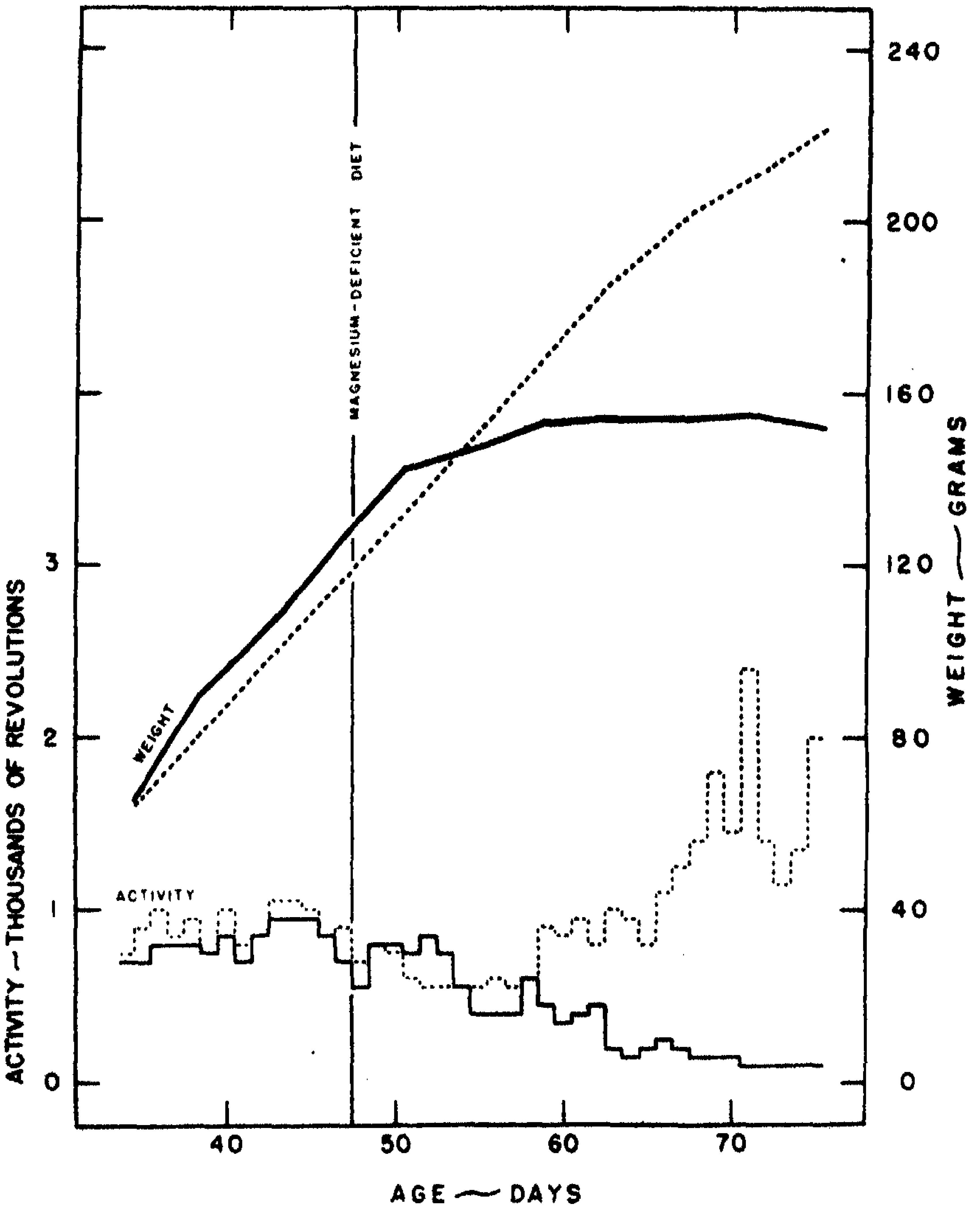


FIGURE 5

Deprivation of magnesium ion. Averaged data from 5 experimental (solid lines) and 5 control animals (broken lines). After the 47th day the magnesium supplement was omitted from the otherwise complete diet of the experimental animals.

have as yet been shown to result in increased running have included deprivation in members of the vitamin B complex.

High running is not, therefore, a reliable sign of well being and optimal performance. It may be a sign of want. When healthy, intact animals are most completely provided with their needs they run minimally. This relation may be used as a criterion of dietary adequacy. Thus on one occasion we have transferred a group of animals back and forth between a complete diet consisting largely of natural foods and an almost wholly synthetic diet containing optimal supplies of all the factors known to be required by rats. Running was consistently higher on the synthetic than on the relatively natural diet, suggesting that the former still lacked factors which the latter contained.

To appreciate the biological significance of running in deprivation it is necessary to realize that it occurs in a behavioral sequence just the reverse of what it is commonly supposed to be. One is used to thinking that an animal first experiences *sensations* of want, then performs appropriate *actions* to satisfy its needs. Such behavior is commonly described as "food-seeking" or "mate-seeking." The futile running of a rat in an activity wheel, however, cannot be characterized in this way; nor can the hunger movements of sleeping animals, or the helpless hunger activity of decorticate animals which make no attempt to take the food placed before them. All these instances show that the primary response to deprivation is activity as such, regardless of sensations or external inducements. If the environment offers the possibility, sensory and cortical mechanisms superimposed upon this primary reaction may guide it into appropriate channels. Otherwise the response remains confined to simple hyperactivity.

In a free environment running should fulfil an important function, since it increases the probability that the animal will encounter what it lacks. In effect it represents a gamble in which the animal's metabolic reserves are staked against the chance of finding its necessities.

The possibility of a successful outcome for the individual, however, is not the only point of this reaction. It probably represents also the behavioral basis of mammalian emigration. This is recognized to be motivated primarily by nutritional need. For the individual emigrant it takes the form of aimless and persistent wandering, without apparent external incentive, and ending almost always in disaster. Its essential biological function is not to rescue the emigrating animal—though this may occur—but to relieve nutritional pressure on the home population. To accomplish this the animal must simply continue to move onward, even though it hasten thereby its own destruction.[15] The running of a deprived animal in an activity wheel provides an accurate model for this behavior.

[*] This research was supported in part by a grant from the Williams-Waterman Fund for the Combat of Dietary Diseases. We were assisted in various portions of the experiments by Miss Anna Cort, Miss Eva Low, and Mrs. Ruth Hubbard, all of Radcliffe College. We are indebted to Distillation Products, Inc., of Rochester, N. Y., for supplies of vitamin E concentrates used in the diets.

[1] Carlson, A. J., *The Control of Hunger in Health and Disease*, Univ. of Chicago Press, Chicago, 1916, Chap. I; Richter, C. P., *Quart. Rev. Biol.*, **2**, 307 (1927).

[2] Wada, T., *Arch. Psychol.*, **8**, no. 57, 1922; Luckhardt, A. B., *Amer. Jour. Physiol.*, **39**, 330 (1915–1916).

[3] Goltz, F., *Arch. ges. Physiol.*, **51**, 570 (1892).

[4] King, J. L., and Connet, H., *Amer. Jour. Physiol.*, **39**, 123 (1915).

[5] Carlson, A. J., *Ibid.*, **31**, 175 (1912–1913); Rogers, F. T., *Ibid.*, **41**, 555 (1916); Pavlov, I. P., *Jour. de Psychol.*, **18**, 273 (1921).

[6] Tsang, Y., *J. Comp. Psychol.*, **26**, 1 (1938).

[7] Wang, G. H., *Comp. Psychol. Monogr.*, **2**, no. 6 (1923); Slonaker, J. R., *Amer. Jour. Physiol.*, **71**, 362 (1925).

[8] Type LC-34, manufactured by the George H. Wahmann Mfg. Co., Baltimore.

[9] Specific deficiency diets were as follows: (*a*) *Thiamin:* the U. S. P. XII thiamin-deficient test ration supplemented with 2 μg. pyridoxine hydrochloride and 10 μg. calcium pantothenate per gram dry food. This was completed on occasion with various amounts of thiamin hydrochloride. (*b*) *Riboflavin:* ration B of Axelrod and Elvehjem (*J. Biol. Chem.*, **140**, 725 (1941)), with its fat content increased to 5 per cent, and supplemented with 10 mgm. mixed tocopherols fed to each animal twice weekly. This was completed when desired with 2 μg. riboflavin per gram of food. (*c*) *Magnesium:* A slight modification of the Mg-low diet of Kruse, *et al.* (*J. Biol. Chem.*, **96**, 519 (1932)), completed when desired with 0.55 per cent $MgSO_4 \cdot 7H_2O$.

[10] Guerrant, N. B., and Dutcher, R. A., *Jour. Nutrition*, **20**, 589 (1940).

[11] Bloomfield, A. L., and Tainter, M. L., *Federation Proc.*, **2**, 75 (1943).

[12] The term *polyneuritis* is used throughout this paper in a restricted sense to mean the stage of thiamin deficiency in which rats exhibit temporary postural disorientation and paralysis after being twirled by the tail (U. S. Pharmacopeia XII, p. 625).

[13] Smith, P. K., and Smith, A. H., *Amer. Jour. Physiol.*, **109**, 99 (1934).

[14] Guerrant, N. B., Dutcher, R. A., and Chornock, F., *Jour. Nutrition*, **17**, 473 (1939).

[15] "The act of mass emigration—is performed by a surplus population. . . . The main incitement—would appear to be want of food. . . and the result in the end is invariably disaster" (Heape, W., *Emigration, Migration and Nomadism*, Heffer and Sons, Cambridge, 1931, p. 21). "It can now be seen that the apparent mystery of where the lemmings are *going to* on their migrations is cleared up when we assume that they are *going away from* the mountains on account of some unfavourable conditions caused by increased density of population" (Elton, C., *Voles, Mice and Lemmings*, Oxford Univ. Press, 1942, p. 215). That man may share to a degree in this response-pattern is suggested by the following account from Herodotus (I, 94) of the invention of athletic games: "In the reign of Atys, son of Manes, there was a great famine throughout all Lydia. The Lydians bore the calamity patiently for some time, but seeing that it did not stop, they set to work to devise remedies for the evil. The device adopted against the famine was to give up one day so entirely to playing as not to feel any want of food, and the next day to eat and stop the games. In this manner they continued for eighteen years. As the affliction lasted and even became more grievous, the king divided the nation in half, and made the two portions draw lots, the one to stay, the other to emigrate from the country."

THE EPIDEMIC CURVE WITH NO ACCESSION OF SUSCEPTIBLES

By Edwin B. Wilson and Jane Worcester

Harvard School of Public Health

Communicated June 22, 1944

In a previous paper[1] we established a higher approximation than $z = z_0 \operatorname{sech}^2 \sqrt{z_0/2}\, T$ to the epidemic curve satisfying Soper's formulation of the problem for such a disease as measles, when recruits to the susceptible population are neglected, namely,

$$T = \frac{1}{6} \log (z/z_0) \pm \frac{1}{3}\sqrt{1 + 18/z_0} \tanh^{-1} \sqrt{1 - z/z_0}, \qquad (1)$$

where T is measured from the peak of the epidemic in units of time equal to the "incubation period" and z_0 is the case rate for new infections at the peak measured in terms of m, the number of susceptibles just sufficient to maintain the disease in a steady state provided an adequate supply of recruits were available. As $z = -dx/dT$ by definition, the number of susceptibles may be obtained by integration, namely,

$$x = \text{const.} + \int z dT = \text{const.} + \int z(dT/dz)dz$$

which, after considerable manipulation, turns out to be

$$x = \text{const.} + \frac{1}{6}(z_0 - z) \pm \frac{1}{3}\sqrt{18 + z_0}\, \sqrt{z_0 - z}, \qquad (2)$$

where the positive sign applies before the peak of the case rate and the negative sign after it. From this form it follows that cases measured relative to m must be,

$$\text{total cases} = 2\sqrt{2z_0}\sqrt{1 + z_0/18} \qquad (3)$$

which may be compared with the result total cases $= 2\sqrt{2z_0}$ for the first approximation $z = z_0 \operatorname{sech}^2 \sqrt{z_0/2}\, T$. As z_0 is a small number compared to 18, the result (3) is not materially different from that obtained before. The ratio of total cases to the case rate at the peak of the epidemic is therefore[2] approximately

$$\frac{\text{total cases}}{\text{peak case rate}} = 2\sqrt{\frac{2}{z_0}} = \frac{2}{B}, \qquad (4)$$

where B is the value obtained in fitting $z = z_0 \operatorname{sech}^2 BT$ and for clear-cut epidemics of measles seems to range from about 0.20 to about 0.60.

To determine the value of x in (2), it is necessary to find the value of the constant of integration. The fundamental equation

$$\frac{dx}{dT} = x\left.\frac{dx}{dT}\right|_{T-1} \quad \text{or} \quad \left.\frac{d\log x}{dT}\right|_{T+1} = \frac{dx}{dT}, \tag{5}$$

when $A = 0$, can be integrated exactly as

$$\log x\,(T+1) = x(T) - k. \tag{6}$$

To effect a comparison of (6) and (2), it is necessary to determine the constant k. We start by expanding $x(T)$ in series with the assumption that the initial conditions are $T = 0$ when $x = 1$, namely,

$$x = 1 + aT + bT^2 + cT^3 + dT^4 + eT^5. \tag{7}$$

Substituting in (6) we find b, c, d, e, in terms of a; the maximum of the case rate curve comes at the value of T such that $d^2x/dT^2 = 0$ and for that value $z = -dx/dT = z_0$. Thus a and the other coefficients may be found in terms of z_0 and finally since $\log x(0) = \log 1 = 0$, $k = x(-1)$ may be found in terms of z_0. The result is[3]

$$\log x(T+1) = x(T) - 1 - z_0 + \frac{z_0^2}{24} - \frac{z_0^3}{180}. \tag{8}$$

This equation may be used to compute $x(T)$ stepwise from any value for $x(T)$ assumed at any time T. If the time T at which $d^2x/dT^2 = 0$ is substituted in (7) we find the value of x at the peak of the epidemic when $z = z_0$, which is precisely the value of the constant in (2). The result is[4]

$$\text{const.} = 1 + \frac{z_0}{2} - \frac{z_0^2}{24} + \frac{59z_0^3}{3840}. \tag{9}$$

Thus, given any value of z_0, we may compare the approximation (2) with the result of a stepwise calculation from (8).

For example, let us return to the hypothesis that $z_0 = 0.3$. Then the relation between the susceptibles x and the case rate z is

$$x = 1.1467 + \frac{1}{6}(0.3 - z) \pm \frac{1}{3}\sqrt{18.3}\,\sqrt{0.3 - z}. \tag{10}$$

Further we have

$$\log x(T+1) = x(T) - 1.2964, \qquad z(T+1) = x(T+1)\,z(T). \tag{11}$$

Starting from the peak at $T = 0$ we have $x(0) = 1.1467$. Then $x(1) = 0.8609$ and $z(1) = 0.2583$. Thus we may compute table 1.

In the table of values the columns for x and z by (11) give the values that are exact in so far as the constant k may be considered as determined from (8) with neglect of powers of z_0 higher than the third; these columns

TABLE 1

VALUES OF x AND z WHEN $z_0 = 0.3$

T	x BY (11)	z BY (11)	x BY (10)	z BY (1)
0	1.1467	0.3000	1.1467	0.3000
+1	0.8609	0.2583	0.8615	0.2580
−1	1.4333	0.2616	1.4331	0.2614
+2	0.6470	0.1671	0.6479	0.1665
−2	1.6563	0.1825	1.6560	0.1821
+3	0.5223	0.0873	0.5239	0.0868
−3	1.8010	0.1102	1.8004	0.1098
+4	0.4611	0.0403	0.4629	0.0400
−4	1.8847	0.0612	1.8836	0.0610
+8	0.4151	0.0013	0.4172	0.0013
−8	1.9725	0.0044	1.9714	0.0043
+∞	0.4136		0.4157	
−∞	1.9790		1.9777	

represent an integration of the fundamental difference equation regulating the epidemic. The columns for x and z by (10) and by (1) are calculated from the approximation (1) which we developed in our previous paper as a good approximation to the integral of the difference equation and from its integral (2). A comparison of the two pairs of columns will therefore show how good the approximation is.

The exact integral (6) enables one to compute the course of a hypothetical epidemic for comparison with the cruder computations by the Soper or Frost finite equations.[5] If we have $m = 1000$ and $S_0 = 2000$ and introduce one case, which will then generate two cases, we have at the start $x(T + 1) = 1.998$ and $x(T) = 2.000$. Then k in (6) is

$$k = 2.000 - \log 1.998 = 1.30785 \qquad (12)$$

and with this value of k, one may use (6) to follow the subsequent course in each step of which the value of x will be cut down so that $S = 1000x$ is integral. The cases may then be obtained by subtracting successive values of x. Table 2 gives the comparison of this calculation with the previous ones. From the result it is seen that the stepwise calculation departs widely from that of Soper but is very much the same as that of Frost (as modified by introducing the law of small numbers).

As a matter of fact we have the equations

$$C_{t+1} = S_t[1 - e^{-rC_t}], \qquad S_{t+1} = S_t - C_{t+1}$$

and if C be eliminated we have

$$S_{t+1} = S_t\, e^{r[S_t - S_{t-1}]}$$

$$\text{or} \qquad \log S_{t+1} - rS_t = \log S_t - rS_{t-1};$$

TABLE 2

COURSE OF HYPOTHETICAL EPIDEMIC $C_0 = 1$, $S_0 = 2000$, $r = 0.001$

GENERATION	BY (6) AND (12)	SOPER	FROST
1	2	2	2
2	4	4	4
3	8	8	8
4	16	16	16
5	31	32	32
6	59	62	61
7	107	116	111
8	180	204	186
9	263	317	268
10	308	393	309
11	271	332	267
12	179	171	173
13	94	59	89
14	43	17	40
15	18	5	17
16	8	1	7
17	3	0	3
18	1	..	1
19	1	..	0
20	0	..	..
Total infected	1596	1739	1594
Residual susceptibles	404	261	406

each of these terms must therefore be the same constant and the relation is therefore equivalent to (6). Hence, although the argument which led to the equations based on considering the epidemic to proceed by discrete generations with the elimination of double contacts was very different from that based on a continuous process and embodied in a differential-difference equation, the final results have been seen to be equivalent.[6]

The characteristics of the curve of case rates (epidemic curve) based upon the general exact solution (8) are not easily obtained, but as table 1 indicates that the values derivable by calculation based upon (11) are so nearly the same as those obtained from the approximation (1), it is of interest to give some of the characteristics of (1), which probably will be a sufficient approximation to the corresponding characteristics of the exact curve. The values

$$\nu_i = \int_{-\infty}^{\infty} T^i z\, dT = \int_{-\infty}^{\infty} T^i z \frac{dT}{dz}\, dz$$

may be obtained by breaking the integral up into the part from $-\infty$ to z_0 and the part from z_0 to ∞ with the appropriate expressions taken from (1). Certain definite integrals are involved, of which some may be evaluated in terms of known constants and others may be had by computation. From the values of ν_i the moments may be computed.

The mean is

$$\frac{1}{3}\log 2 - \frac{1}{2} = -\ 0.26895,$$

$$\mu_2 = \frac{\pi^2}{6z_0} + 0.07070, \qquad \mu_3 = -\frac{1.2020}{z_0} - 0.10424.$$

From this it is seen that the mean (referred to the mode) is negative and independent of z_0. Except for the small term 0.07070, the square of the standard deviation is the same as that for the first approximation $z = z_0 \operatorname{sech}^2 \sqrt{z_0/2}T$. The negative value of μ_3 shows that the skewness is negative (the epidemic curve rises more slowly than it falls) and if the small additive terms be neglected it is $\kappa = \mu_3/\mu_2^{3/2} = -0.57z_0^{1/2}$.

The skewness is not only negative but large. Brownlee noted that observed epidemics were remarkably symmetrical and that such skewness as they had was much more likely to be positive than negative; he inferred therefrom that the want of persons liable to infection is not the cause of the decay of the epidemic.[7] Even though the want of persons liable to infection is not a sufficient cause to explain the slope of epidemic curves, it does seem that for immunizing diseases like measles it must be a necessary cause in any ultimate explanation. It is therefore of interest to have (1) an exact integral of the differential-difference equation which regulates that part of the phenomenon that derives from the law of mass action and (2) a satisfactorily close approximation to that integral in manageable form.

---

[1] These PROCEEDINGS, **30**, 37–44 (1944).

[2] If we take the four epidemics used in figure 1 of the previous paper, we have:

EPIDEMIC	TOTAL CASES	PEAK CASE RATE	RATIO	$2/B$
Vancouver, B. C., 1931–1932	3898	1150	3.4	3.6
Berkeley, Calif., 1938–1939	2073	625	4.9	4.8
Providence, R. I., 1934–1935	5267	1000	5.3	4.8
Oakland, Calif., 1933–1934	4889	625	7.8	8.0

The Vancouver epidemic was extremely explosive. The values of the peak case rate are roughly estimated. The agreement between the last two columns is reasonably satisfactory.

[3] The value of k goes only up to terms in z_0^3 although x was expanded to T^5. There is "loss of places" in an algebraic sense in carrying out the work.

[4] In these PROCEEDINGS, **29**, 43–48 (1943), the expression $z = z_0 \operatorname{sech}^2 \sqrt{z_0/2}\,T$, or its equivalent, was integrated to obtain $x = 1 - \sqrt{2z_0} \tanh \sqrt{z_0/2}\,T$, or equivalent, on the assumption that at the peak when $T = 0$ the value of x would be 1. This was in fact one of Soper's assumptions in that he took, in our notation, $1 + du/dT = x$ as one of the starting points of his developments, and this makes $x = 1$ when u is a maximum. It may be that to the order of approximation which Soper uses, it is satisfactory to take $x = 1$ at maximum; but if we go back to the equation from which he starts, viz., $z(T) = x(T)z(T-1)$ we find that

$$x = \frac{z}{z - \frac{dz}{dT} + \frac{1}{2}\frac{d^2z}{dT^2}} = 1 + \frac{1}{z}\frac{dz}{dT} - \frac{1}{2z}\frac{d^2z}{dT^2} + \frac{1}{z^2}\left(\frac{dz}{dT}\right)^2$$

so that if $dz/dT = 0$ when $z = z_0$ we obtain

$$x = 1 - \frac{1}{2z_0}\left.\frac{d^2z}{dT^2}\right|_0 \text{ and from } d^2u\,\frac{d^2u}{dT^2} = -e^u = -z$$

to which he comes, after his various approximations, we find $x = 1 + z_0/2$ as presumably a better result, and this checks with (9). This would make

$$x = 1 + \frac{1}{2}z_0 - \sqrt{2z_0}\tanh\sqrt{z_0/2}\,T$$

with the initial and final values of x, $x_0 = 1 + \frac{1}{2}z_0 + \sqrt{2z_0}$, $x_E = 1 + \frac{1}{2}z_0 - \sqrt{2z_0}$ and with the relationship $x_0 + x_E = 2 + z_0$ instead of $x_0 + x_E = 2$.

[5] These PROCEEDINGS, **28**, 361–367 (1942).

[6] The slight numerical differences between the second and fourth columns in table 2 are due to differences in the rejection of fractional cases.

[7] Brownlee, J., *Proc. Roy. Soc. Med., London*, **2**, Pt. 2, 243–258 (1909). See page 244. This is quoted (on page 206) by Sir Ronald Ross, *Proc. Roy. Soc. London*, **A92**, 204–230 (1915), and by C. O. Stallybrass on p. 510 of his *Principles of Epidemiology*, Chap. 14, 1931.

INTEGRAL FORMULAS FOR THE CHARACTERISTIC CLASSES OF SPHERE BUNDLES

BY SHIING-SHEN CHERN

INSTITUTE FOR ADVANCED STUDY, PRINCETON, NEW JERSEY AND TSING HUA UNIVERSITY, KUNMING, CHINA

Communicated July 13, 1944

The notion of sphere bundles[1] of H. Whitney has proved to be very fruitful in the applications of topology. So far in this theory an important rôle has been played by a class of topological invariants in the base space, called the characteristic cohomology classes and which can be regarded as generalizations of the Euler-Poincaré characteristic. By assuming the base space to be a differentiable manifold with a Riemannian metric, on which the bundles of vectors and p-vectors are defined in a natural way, we shall give in this note various relations which exist between these topological invariants and certain differential invariants of the Riemannian manifold. In fact, the relations are such that the former are expressed as integrals of the latter. In the formulae the generalized Gauss-Bonnet formula due to Allendoerfer and Weil[2] is included as a particular case.

1. At the basis of our considerations is a closed orientable n-dimensional Riemannian manifold of class C^m, $m \geqq 4$, which we denote by R^n. In R^n

we consider the frames $P e_1 \ldots e_n$, each of which consists of a point P and n mutually perpendicular unit vectors $e_1, \ldots, e_n$ through P, arranged in an order coherent with the orientation of R^n. Using the notations of E. Cartan,[8] the parallelism of Levi-Civita is defined by equations of the form

$$dP = \omega_i e_i, \qquad de_i = \omega_{ij} e_j, \qquad \omega_{ij} + \omega_{ji} = 0, \tag{1}$$

where the ω's are Pfaffian forms. In terms of these Pfaffian forms the equations of structure of the space are

$$d\omega_i = \omega_j \omega_{ji}, \qquad d\omega_{ij} = -\omega_{ik}\omega_{jk} + \Omega_{ij}, \tag{2}$$

and the identities of Bianchi are

$$\omega_j \Omega_{ji} = 0, \qquad d\Omega_{ij} = -\omega_{ik}\Omega_{jk} + \omega_{jk}\Omega_{ik}. \tag{3}$$

It is to be mentioned that Ω_{ij} are exterior quadratic differential forms which give the curvature properties of R^n.

Consider now the tangent sphere bundle of R^n in the sense of Whitney and, in particular, the problem of defining over R^n a field of p-vectors, $1 \leqq p \leqq n$, where by a p-vector is meant an ordered set of p vectors (and not the figure determined by the Grassmannian coördinates). As is well known, the characteristic classes arise from the points of R^n at which the p vectors are linearly dependent. Topologically there is no loss of generality to assume the p vectors, when linearly independent, to be mutually perpendicular unit vectors.

We are thus led to consider in R^n the figures formed by a point P and an ordered set of p mutually perpendicular unit vectors through P. With P fixed, all these figures form, according to a natural topology, a finite manifold $V_{n,p}$ of dimension $p(2n - p - 1)/2$. The Betti group $B^{n-p}(V_{n,p})$ of dimension $n - p$ (with integer coefficients) of $V_{n,p}$ is an infinite cyclic group or a cyclic group of order two, according as one of the following conditions is satisfied or not: $p = 1$ or $n - p$ is even. We consider the sphere bundle $\mathfrak{S}(p)$ with R^n as the base space and the $V_{n,p}$ as fibres and proceed to study the relations between the characteristic classes arising therefrom and certain differential invariants of R^n.

2. With our problem formulated in this way, it is convenient to adopt the language of the method of moving frames. To a figure of the above description we attach a frame $Pe_1 \ldots e_n$ such that P is the origin and $e_{n-p+1}, \ldots, e_n$ the p vectors of the figure in question. With the intervention of the family of frames thus attached, Pfaffian forms ω_i, ω_{ij} and differential invariants can be constructed from (1, 2). From them it is possible to construct exterior differential forms of higher degree. Whether they belong to $\mathfrak{S}(p)$ or not can be decided by the following simple criterion: *A differential form* Π *belongs to* $\mathfrak{S}(p)$ *if and only if its exterior derivative* $d\Pi$ *is a form in* ω_i, ω_{iA} $A = n - p + 1, \ldots, n$, *only*. A differential form belonging to $\mathfrak{S}(p)$ is called intrinsic.

3. For our considerations we shall suppose either $p = 1$ or $n - p$ be even. The essential step consists of establishing a formula for differential forms which we shall reproduce here for the case that $n - p$ is even. We consider the following intrinsic differential forms

$$(2k;\ h_1, \ldots, h_p) = \epsilon_{\alpha_1 \ldots \alpha_{n-p}} \Omega_{\alpha_1\alpha_2} \ldots \Omega_{\alpha_{2k-1}\alpha_{2k}} \omega_{\alpha_{2k+1}\, n-p+1} \ldots$$
$$\omega_{\alpha_{2k+h_1}\, n-p+1}\, \omega_{\alpha_{2k+h_1+1}\, n-p+2} \ldots \omega_{\alpha_{2k+h_1+h_2}\, n-p+2} \ldots$$
$$\omega_{\alpha_{2k+h_1+\ldots+h_{p-1}+1}\, n} \ldots \omega_{\alpha_{2k+h_1+\ldots+h_p}\, n} \qquad (4)$$

where each index α runs from 1 to $n - p = 2k + h_1 + \ldots + h_p$ and where $\epsilon_{\alpha_1 \ldots \alpha_{n-p}}$ denotes the Kronecker symbol which is $+1$ or -1 according as $\epsilon_{\alpha_1}, \ldots, \alpha_{n-p}$ form an even or odd permutation of $1, \ldots, n - p$, and is otherwise zero. We put also

$$\Phi_k = \sum \frac{2.4 \ldots (n-p-2k)}{(2.4 \ldots 2q_1) \ldots (2.4 \ldots 2q_p)} (2k;\ 2q_1, \ldots, 2q_p) \qquad (5)$$

where the summation is extended over all partitions of $^1/_2(n - p) - k$ as a sum of $q_1, \ldots, q_p$. Further let

$$a_k = (-1)^k \frac{(n-p-2k+2) \ldots (n-p)}{2.4 \ldots 2k}, \ k \geqq 1, \qquad (6)$$
$$a_0 = 1.$$

Then we have the formula

$$d\left\{ \sum_{k=0}^{1/2(n-p)} a_k \Phi_k \right\} = 0. \qquad (7)$$

The proof of this formula can be carried out in a purely algebraic way.

4. Suppose we consider the p-vectors having the same origin, i.e., those belonging to the same $V_{n,p}$. Under this condition we have evidently

$$\omega_1 = \ldots = \omega_n = 0,$$

and formula (7) becomes

$$d\Phi_0 = 0$$

This shows that Φ_0 is a closed intrinsic form in $V_{n,p}$. Noticing that $B^{n-p}(V_{n,p})$ is an infinite cyclic group, it follows from the well-known results of de Rham[4] that the value of the integral

$$\int_Z \Phi_0$$

where Z is an $(n - p)$-dimensional cycle of $V_{n,p}$, will give the homology class to which Z belongs.

The theory of p-fields in R^n leads, when $n - p$ is even, to an integral characteristic class W^{n-p+1}. To derive from it a numerical topological invariant we consider the scalar product $W^{n-p+1}M^{n-p+1}$, where M^{n-p+1} is a submanifold of dimension $n - p + 1$ in R^n, closed or with boundary. When M^{n-p+1} is closed, we derive from (7) and the above remark that

$$W^{n-p+1}.M^{n-p+1} = 0,$$

which corresponds to the known fact that W^{n-p+1} is of finite order.

If, however, M^{n-p+1} has a boundary B, we shall get in this way a formula which links the integral of a differential invariant over B with a topological invariant of M^{n-p+1}. For definiteness we suppose $p = 2$ (so that n is even). B is of dimension $n - 2$ and we suppose for simplicity that B possesses a tangent $(n - 2)$-plane at each of its points. Both M^{n-1} and B are supposed to be orientable, so that M^{n-1} is two-sided in R^n and B two-sided in M^{n-1}. At each point of B we take as e_n the outside (or inside) normal of M^{n-1} and as e_{n-1} the outside (or inside) normal of B in the tangent hyperplane of M^{n-1}. We have thus defined over B a field $\mathfrak{F}$ of two vectors and the form

$$\Pi = \sum_{k=0}^{1/2(n-2)} a_k\Phi_k,$$

being intrinsic in $\mathfrak{S}(2)$, is defined on B. Then we have

$$\int_B \Pi = c.M^{n-1}.LL^{n-1}$$

where c is a numerical constant. This formula is the analog of the Gauss-Bonnet formula for a polyhedron. The corresponding formulation for a general p is without difficulty.

5. As we have shown in another occasion,[6] these discussions lead for $p = 1$ to a proof of the Gauss-Bonnet formula by which it is not necessary to imbed the cells of R^n in a Euclidean space of sufficiently high dimension. There the formula corresponding to (7) is of the form

$$d\Pi = \Omega,$$

where $\Omega = 0$ when n is odd. The differential form Π, of degree $n - 1$, is defined in the space of vectors and has some interest in itself. It is defined, whenever an orientable submanifold M^r, of dimension r, is given in R^n. For, supposing that M^r has a tangent r-plane at each of its points, we attach to a point of M^r all the unit normal vectors at that point if $r \neq n - 1$ and the outside (or inside) unit normal vector at that point if $r = n - 1$. In this way the set of unit vectors attached to the points of M^r forms, according to the natural topology, a manifold of dimension $n - 1$. This process can also be carried out for M^r which ceases to have tangent r-

planes at certain points. Then we attach to each of these points the unit normal vectors belonging to the corresponding dual angle. By such a convention the differential form II is defined over M^r and generalizes the notion of the element of geodesic curvature for $n = 1$. We can prove that, if r is even,

$$\int_{M^r} \mathrm{II} = c\chi(M^r),$$

where $\chi(M^r)$ stands for the Euler-Poincaré characteristic of M^r and c is a constant. The proof of this formula is somewhat complicated.

6. It is evident that the same considerations can be carried out for the *normal sphere bundles* of a *submanifold imbedded in a differentiable manifold*. To avoid duplication we shall not give the results here. We only mention that the corresponding differential invariants are those which give the so-called Gaussian torsion of the submanifold.

7. Finally, we wish to make the following remark. The process employed here bears a close resemblance to that used by Gysin[6] in his work on the homology theory of fibre spaces. Our differential forms take the place of cochains and we have taken the group of real numbers as the coefficient group of the homology theory. While we thus lose such topological invariants like torsion, we are compensated by the fact that the differential forms are intrinsic in the sense of differential geometry and have thus arrived at results which connect differential invariants with topological invariants.

[1] Whitney, H., "On the Topology of Differentiable Manifolds," *Lectures in Topology* Michigan, 1941, pp. 101–141. Also: Stiefel, E., "Richtungsfelder und Fernparallelismus in n-dimensionalen Mannigfaltigkeiten," *Comm. Math. Helv.*, **8**, 305–353 (1936).

[2] Allendoerfer, C. B., and Weil, André, "The Gauss-Bonnet Theorem for Riemannian Polyhedra," *Trans. Amer. Math. Soc.*, **53**, 101–129 (1943).

[3] Cartan, E., *Leçons sur la géometrie des espaces de Riemann*, Paris, 1928.

[4] de Rahm, G., "Sur l'analysis situs des variétés a n dimensions," *J. Math. Pures Appl.*, **10** (9), 115–200 (1931).

[5] Chern, S., "A Simple Intrinsic Proof of the Gauss-Bonnet Formula for Closed Riemannian Manifolds," to appear in *Annals of Math.*

[6] Gysin, W., "Zur Homologietheorie der Abbildungen und Faserungen von Mannigfaltigkeiten," *Comm. Math. Helv.*, **14**, 61–122 (1941–1942).

ON GROUPS OF AUTOMORPHISM OF LIE GROUPS

By C. Chevalley

Department of Mathematics, Princeton University

Communicated July 31, 1944

Let $\mathfrak{G}$ be a simply connected Lie group, and let $\mathfrak{g}$ be the Lie algebra of $\mathfrak{G}$. Then the group of automorphisms $\mathfrak{A}$ of $\mathfrak{G}$ is isomorphic to a linear group $\mathfrak{A}^*$ operating on the vector space $\mathfrak{g}$. The infinitesimal transformations of $\mathfrak{A}^*$ are the linear mappings D of $\mathfrak{g}$ into itself which have the property that

$$D([X, Y]) = [DX, Y] + [X, DY] \qquad (1)$$

holds for any X and Y in $\mathfrak{g}$. These mappings are called the *derivations* of the Lie algebra $\mathfrak{g}$. They form a Lie algebra which is called the *algebra of derivations* of $\mathfrak{g}$. We shall denote this algebra by $\mathfrak{D}(\mathfrak{g})$.

The inner automorphisms of $\mathfrak{G}$ form a distinguished subgroup $\mathfrak{A}_0$ of $\mathfrak{A}$. The above-mentioned isomorphism of $\mathfrak{A}$ with $\mathfrak{A}^*$ maps $\mathfrak{A}_0$ onto the adjoint group $\mathfrak{A}_0^*$ of $\mathfrak{G}$. The infinitesimal transformations of $\mathfrak{A}_0^*$ are the derivations D_Z of $\mathfrak{g}$ which can be defined by $D_Z(X) = [X, Z]$ (for some fixed $Z \epsilon \mathfrak{g}$). They are called the *inner derivations* of $\mathfrak{g}$. They form an ideal in $\mathfrak{D}(\mathfrak{g})$, and this ideal is isomorphic to the factor algebra $\mathfrak{g}/\mathfrak{z}$, where $\mathfrak{z}$ is the center of $\mathfrak{g}$. We say that $\mathfrak{g}$ is a *complete algebra* when every derivation of $\mathfrak{g}$ is an inner derivation. Cartan has proved that every semi-simple Lie algebra is complete. Zassenhaus has proved that, if $\mathfrak{g}$ is a Lie algebra whose center reduces to $\{O\}$ and which coincides with its derived algebra, then $\mathfrak{D}(\mathfrak{g})$ is a complete algebra.

Starting with an arbitrary Lie algebra $\mathfrak{g}$, we can define inductively a sequence of Lie algebras $\mathfrak{D}_n(\mathfrak{g})$ in the following way: $\mathfrak{D}_0(\mathfrak{g}) = \mathfrak{g}$; if $\mathfrak{D}_n(\mathfrak{g})$ has been defined, $\mathfrak{D}_{n+1}(\mathfrak{g})$ is $\mathfrak{D}(\mathfrak{D}_n(\mathfrak{g}))$. The object of this note is to announce the following result:

THEOREM. *Let $\mathfrak{g}$ be a Lie algebra over a field of characteristic 0, and assume that the center of $\mathfrak{g}$ contains only 0. Then there exists an index n such that $\mathfrak{D}_n(\mathfrak{g})$ is complete, and therefore all algebras $\mathfrak{D}_n(\mathfrak{g})$ are isomorphic with each other from a certain rank on.*

Only broad outlines of the proof can be given here. It is clearly sufficient to prove the theorem for Lie algebras over the field of complex numbers. Under this condition, $\mathfrak{D}(\mathfrak{g})$ is the Lie algebra of an *algebraic* linear group. From this fact, one can deduce that $\mathfrak{D}(\mathfrak{g})$ may be represented as the direct sum of three subalgebras $\mathfrak{s}$, $\mathfrak{a}$ and $\mathfrak{n}$ with the following properties: $\mathfrak{s}$ is semi-simple; $\mathfrak{a}$ is abelian; $\mathfrak{n}$ is the maximal nilpotent ideal; the matrices in $\mathfrak{a}$ can be simultaneously reduced to the diagonal form, and they commute with the elements of $\mathfrak{s}$.

Let $\mathfrak{n}_0$ be the subalgebra of $\mathfrak{n}$ generated by the elements of the form $[A,$

$N]$ or $[A, S]$, with $A \in \mathfrak{a}$, $N \in \mathfrak{n}$ $S \in \mathfrak{f}$, and set $\mathfrak{d}_0 = \mathfrak{f} + \mathfrak{a} + \mathfrak{n}_0$. Then $\mathfrak{d}_0$ is an ideal in $\mathfrak{D}(\mathfrak{g})$, and the centralisator of $\mathfrak{d}_0$ in $\mathfrak{D}(\mathfrak{g})$ contains only O. Let $J(\mathfrak{d}_0)$ be the ideal formed in $\mathfrak{D}(\mathfrak{d}_0)$ by the inner derivations of $\mathfrak{d}_0$. By a derivation of $J(\mathfrak{d}_0)$ into $\mathfrak{D}(\mathfrak{d}_0)$ is meant a linear mapping of $J(\mathfrak{d}_0)$ into $\mathfrak{D}(\mathfrak{d}_0)$ which satisfies condition (1) above for any X and Y in $J(\mathfrak{d}_0)$. It can be proved that any such derivation maps $J(\mathfrak{d}_0)$ into itself and is therefore induced by some inner derivation of $\mathfrak{D}(\mathfrak{d}_0)$. From this, one can deduce that, $\mathfrak{h}$ being any subalgebra of $\mathfrak{D}(\mathfrak{d}_0)$ containing $J(\mathfrak{d}_0)$, $\mathfrak{D}(\mathfrak{h})$ is isomorphic to some subalgebra of $\mathfrak{D}(\mathfrak{d}_0)$ containing $\mathfrak{h}$, from which it follows immediately that $\mathfrak{D}_n(\mathfrak{h})$ is complete for sufficiently large n. On the other hand, one can see that $\mathfrak{D}(\mathfrak{g})$ itself is isomorphic to a subalgebra of $\mathfrak{D}(\mathfrak{d}_0)$ containing $J(\mathfrak{d}_0)$, whence the result to be proved.

The restriction that the center of $\mathfrak{g}$ be $\{O\}$ cannot be removed from the statement of our Theorem, because examples show that the following two situations may arise by proper choices of $\mathfrak{g}$: I, the dimension of $\mathfrak{D}_n(\mathfrak{g})$ increases indefinitely with n; II, the algebras $\mathfrak{D}_n(\mathfrak{g})$ are all isomorphic with each other for $n \geqslant 1$, but none of them is complete. Nevertheless, there are large classes of nilpotent Lie algebras $\mathfrak{g}$ for which $\mathfrak{D}_n(\mathfrak{g})$ ultimately becomes complete. Let, for instance, L_q be the free Lie ring with q generators, and define $L_q^{(r)}$ by induction on r in the following way: $L_q^{(0)} = L_q$; $L_q^{(r+1)}$ is the set of linear combinations of elements of the form $[X, Y]$, with $X \in L_q$, $Y \in L_q^{(r)}$. Then $L_q^{(r)}$ is an ideal in L_q and $L_q/L_q^{(r)}$ is a nilpotent algebra $\mathfrak{g}(q, r)$. It can be proved that $\mathfrak{D}_2(\mathfrak{g}(q, r))$ is complete provided $q > 1, r > 1$.

On the other hand, it follows immediately from our theorem that, if $\mathfrak{g}$ coincides with its derived algebra, then $\mathfrak{D}_n(\mathfrak{g})$ is complete for n sufficiently large. In fact, it is easy to see that the center of $\mathfrak{D}_1(\mathfrak{g})$ contains only O.

GROUPS CONTAINING LESS THAN TWENTY-EIGHT NON-INVARIANT OPERATORS

By G. A. Miller

Department of Mathematics, University of Illinois

Communicated July 19, 1944

In two preceding numbers of these PROCEEDINGS (**30**, 25–28 (1944), **30**, 114–117 (1944)) various methods were noted for the determination of the possible groups when the number of the non-invariant operators is given. The object of the present article is to give a few additional such methods and to illustrate these methods by means of the determination of all the groups which separately involve less than twenty-eight non-invariant operators. It may be noted that in the early determination of permutation groups and abstract groups A. L. Cauchy (1789–1857), A. Cayley (1821–1895) and others proceeded similarly with a view to introducing their readers to the subjects which were later extensively developed by others by means of improved methods.

A useful theorem in the said determination may be expressed as follows: *If a group of order* $m + n$ *contains* n *invariant operators and* m *non-invariant ones then* n *is a divisor of* m *and cannot exceed* $m/2$. This theorem is evident but it seems desirable to state it explicitly in order to simplify its applications in what follows. It may be desirable to recall here the theorem that if a group contains a prime number p of non-invariant operators and all of its other operators are invariant then it contains exactly one invariant operator and therefore its order is $p + 1$. As illustrative examples we note the non-cyclic group of order 6, the tetrahedral group, and the simple group of order 60, which is the simple group of smallest composite order.

Since all the groups which separately contain less than twenty-four non-invariant operators can readily be determined by means of the noted theorems we proceed to consider here, with some details, all the groups which contain separately exactly twenty-four non-invariant operators. It is obvious that the order of such a group is either 27 or 32, and if it is 27 the group must be one of the two non-abelian groups of this order. It is therefore only necessary to consider groups of order 32 in determining the remaining groups which separately contain twenty-four non-invariant operators. Moreover, the only groups of order 32 which require consideration in this connection are those which contain separately just eight invariant operators and hence have for their central quotient group the non-cyclic group of order 4, since the central quotient group of a group cannot be cyclic.

It therefore results that all the groups of order 32 which we need to consider here contain three abelian subgroups of index 2 whose cross-cut is the central of the group G. Moreover, the commutator subgroup of each of these groups is in the central of G and is of order 2. It will be convenient to consider separately the possible groups when the central is a given group beginning with the case when this central is cyclic. In this case G may contain a cyclic group of order 16 and if it has this property two of the three abelian subgroups of index 2 in G are cyclic while the third is of type 3, 1. There is obviously one and only one such G. When the central of G is cyclic but G does not involve any operator of order 16 the three abelian subgroups of G are all of type 3, 1 and there is clearly one and only one such G. We have now determined the two possible groups when the central of G is cyclic.

The central of G in the remaining groups is either of type 2, 1 or of type

1, 1, 1. We begin with the former. of these two possible cases. In this case the commutator subgroup of G is either the square of the operators of order 4 contained in the central or it is one of the other two operators of order 2 in the central. We begin with the case when the central is of type 2, 1 and when a commutator is the square of an operator of order 4 contained in the central. If such a G contains operators of order 8 there are two possible groups. In both of these groups two of the three abelian subgroups of index 2 are of type 3, 1. In one of them the third subgroup of this index is of type 2, 1, 1 while in the other it is of type 2, 2. To complete the determination of the groups which involve operators of order 8 we have yet to consider the case when the commutator of order 2 is not the square of an operator of order 4 contained in the central of G.

In this case there are again two groups in which two of the three abelian subgroups of index 2 are of type 3, 1. In one of these the third abelian subgroup of index 2 is again of type 2, 1, 1 while in the other it is of type 2, 2. In these groups the cyclic subgroups of order 8 are non-invariant while in each of the preceding four groups the corresponding cyclic subgroups are invariant. It remains to determine the groups which involve no operator whose order exceeds 4 and in which the central is either of type 2, 1 or of type 1, 1, 1. If in the former case G involves an abelian subgroup of index 2 and of type 2, 2 it may first be assumed that the commutator subgroup of G is generated by an operator of order 4 contained in the central of G. In this case the remaining operators of G must include an operator of order 2. If this were not the case all the remaining operators of G would be of order 4 and could not have for their squares a commutator of G.

If s_1 is such an operator and s_2 is an operator of order 4 in the given subgroup of type 2, 2 which is not in the central then $(s_1s_2)^2 = s_1^2s_2^2$ into the commutator of order 2 in G. It therefore results that G would involve an operator of order 2 which is not contained in the given subgroup of type 2, 2. There is clearly one and only one such group. One of its three abelian subgroups of index 2 is of type 2, 1, 1 while the other two are both of type 2, 2. When the commutator of order 2 in G is not the square of an operator of order 4 in the central of G but G still involves an abelian subgroup of type 2, 2 there are two additional groups. In one of these all of the remaining operators of G are of order 4 while in the other only one-half of these operators have this property. We have now considered all the groups which contain no operator of order 8 and in which one of the abelian subgroups of index 2 is of type 2, 2.

When G contains a central of type 2, 1 but no abelian subgroup of type 2, 2 and no operator of order 8 its three abelian subgroups of index 2 are all of type 2, 1, 1 and the commutator subgroup is generated by an operator of order 4 contained in the central of G. There is one and only one such group. It remains to consider the case when the central is of type 1, 1, 1.

In this case at least one of the abelian subgroups of index 2 in *G* must be of type 2, 1, 1. If the commutator subgroup is generated by an operator of order 4 in this subgroup of index 2 all the operators of this subgroup are transformed into their inverses and there are three such distinct groups. In one of these all of the remaining operators are of order 2; in another all of these operators have the same square as the operators of order 4 in the given subgroup of index 2, while in the third all of these operators have for their squares a different operator of order 2 contained in this subgroup.

Finally, when the commutator subgroup of *G* is a different subgroup of order 2 contained in the central there are two additional groups. One of these contains operators of order 2 which are not found in the given subgroup of index 2 while in the other all of these remaining operators are of order 4. We have now determined the seventeen groups which separately contain twenty-four non-invariant operators while all of their other operators are invariant. Two of these groups are of order 27 while each of the remaining fifteen groups is of order 32. These fifteen groups could also be determined by means of the fifty-one known groups of order 32, but the direct determination by means of the fact that each of them contains exactly eight invariant operators seems to be an easier method and illustrates the advantage of considering groups from the standpoint of the number of their non-invariant operators.

If a group contains exactly twenty-five non-invariant operators it may be the dihedral group of order 26 or the direct product of the group of order 5 and the non-abelian group of order 6. As it could obviously be no other group it is a very simple problem to determine all the groups which separately contain exactly twenty-five non-invariant operators. There are also just two groups which separately contain exactly twenty-six non-invariant operators. Both of these groups contain the cyclic group of order 14 and transform all of its operators into their inverses. In one of these groups all of the remaining operators are of order 2 while in the other all of these operators are of order 4 and have a common square. Hence the possible groups which separately contain twenty-six non-invariant operators are also very elementary.

When a group contains exactly twenty-seven non-invariant operators it may be the direct product of the dihedral group of order 10 and the group of order 3. It could not be another group of order 30 since every group of this order contains an invariant subgroup of order 5. Its order could not exceed 36 since the order of a group cannot exceed the number of its non-invariant operators multiplied by 4/3. Its order could not be 36 since its central would then be of order 9 and hence its commutator subgroup could not be of order 2 as it must be in every group of order g in which the number of the non-invariant operators is $3\,g/4$. Hence it results that all the groups which contain less than twenty-eight non-invariant operators have been determined. When the number of the non-invariant operators of a group is twenty-eight the group may contain a central of order 4 and be of order 32. The number of possible groups in this case seems to be large.

A GROUP-THEORETIC CHARACTERIZATION OF THE GENERAL PROJECTIVE COLLINEATION GROUP*

BY N. S. MENDELSOHN

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF TORONTO

Communicated July 11, 1944

The success which investigators in the nineteenth century achieved in the study of geometry through the application of group-theoretic ideas may be credited to the celebrated Erlanger Program of Klein. Briefly stated, the main idea of this Program was to consider geometry as a body of invariant theory of transformation groups. The problem immediately arose as to how any given geometry could be generated in this way. To answer such a question one could assume an abstract group and by introducing suitable definitions try to set up a geometry of the desired type. This has been done for plane Euclidean geometry by Thomsen[1] and non-Euclidean geometries by Hjelmslev.[2] The purpose of the present paper is to summarize a solution of the problem in the case of projective geometry.

We shall start with an abstract group and shall define certain transformations in it by means of its inner automorphisms. These inner automorphisms are used to define a projective geometry which has a projective collineation group isomorphic to the original abstract group. As a corollary, we obtain a very simple representation of the complete group of automorphisms of the projective collineation group, without having recourse to coördinates. By a simple specialization of the projective group one can also obtain a representation of Euclidean geometry. We consider here mainly the two-dimensional case; extension to n-dimensions, while by no means immediate, is not difficult.

Let us denote the normalizer of the subgroup P of an abstract group G by the notation $N(P)$. Two subgroups A and B of G will be said to be *strongly* conjugate in G if $B = g^{-1}Ag$, where g can be chosen to be of finite order. We state the following axioms.

AXIOM 1. *Let G be a group generated by two complete sets of strongly conjugate subgroups $\{P\}$ and $\{L\}$ such that the only element common to all subgroups of both sets is the identity.*

DEFINITION: *A subgroup of class $\{P\}$ will be called a point, one of class $\{L\}$ a line.*

AXIOM 2. *If T is a point or line $NN(T) = N(T)$.*

AXIOM 3. *If $A \subset N(D)$, there exist two other subgroups B, C conjugate to A and such that $(B, C) \subset N(D)$. Of the two groups A and D, one is of class $\{P\}$ and the other of class $\{L\}$.*

AXIOM 4. (*a*) *For any distinct points P and Q, $(P, Q) = N(L)$ for some line L.* (*b*) *For any distinct lines L and M, $(L, M) = N(P)$ for some point P.*

DEFINITION: *A point P is said to be on a line L if $P \subset N(L)$. A line L is on a point P if $L \subset N(P)$. This symmetry of the relation on has to be proved.*

AXIOM 5. *If A, B, C are three distinct points (lines) on the line (point) D, then $N(A) \cap N(B) \cap N(C) = D$.*

AXIOM 6. *Let P and L be any point and line, respectively. Then*

$$G = N(L) + \sum N(L)g,$$

and either $P \subset N(L)$ or for each coset $N(L)g$ such that P not-$\subset g^{-1}N(L)g$, the set $P \cap N(L)g$ is non-null.

Remarks Concerning the Axioms.—As they stand the axioms are not sufficient to define projective geometry, but this can be accomplished by strengthening Axiom 6. The Axioms are consistent since they can be shown to hold in any finite projective geometry. By means of examples, each of Axioms 5, 4, 3, and 1 and 2 together can be shown to be independent of the remaining ones. In Axiom 1 it was found necessary to restrict our sets $\{P\}$ and $\{L\}$ by the condition of "strong" conjugacy. The reason for this is that we are not limiting ourselves to finite groups. It is possible to give examples of infinite groups from which an infinite sequence of subgroups $A_1, A_2, \ldots A_n, \ldots$ can be chosen, such that A_i is conjugate to A_1 but for which $A_i \subset A_{i+1}$ strictly. The use of strong conjugacy eliminates this possibility for the groups $\{P\}$ and $\{L\}$.

That the points P and the lines L form a projective geometry is proved by the following sequence of theorems and definitions:

A point is on a line if and only if the line is on the point.

The group of a point or line is uniquely determined by its normalizer.

Through two points there passes one and only one line, and, dually, on two lines there is only one point.

The condition that points A, B, C, are collinear is that $C \subset (A, B)$, and this is equivalent to the condition $A \subset (B, C)$ or $B \subset (A, C)$.

Any element of G is defined to be a projective collineation. If h is an element of G, then h is said to carry the point A into the point B if $h^{-1}A\,h \subset B$.

Any projective collineation carries collinear points into collinear points and concurrent lines into concurrent lines.

A projective collineation t for which $t^{-1}P\,t \subset P$ is said to keep the point P fixed. The set of all projective collineations which keep P fixed form the group $N(P)$ and dually for lines. The group P consists of all projective collineations, each of which keeps every line through P fixed and dually. Perspectivity is defined by means of Axiom 6. We show that corresponding to any two lines L_1 and L_2 and any point O not on either line, there exist collineations t which keep O fixed and which carry L_1 into L_2. The class of all such elements t, is called *perspectivity* of the line L_1 on L_2 from the center O. Dually, perspectivity from a line may be defined.

It follows now that a necessary and sufficient condition for a projectivity t between two lines L_1 and L_2 to be a perspectivity is that t keep fixed the common point of L_1 and L_2. This is equivalent[3] to Pappus' Theorem from which Desargues' Theorem follows in the plane. By sufficiently weakening our axioms one could obtain a geometry for which Desargues' Theorem is not valid but for which the weaker condition, namely, the unique determination of the fourth harmonic point, holds. Such geometries have, in fact, been considered from the geometric point of view, but a study of their groups would also be of interest.

If A, B, C, D and A', B', C', D' are two sets of four points, no three of which in either set are collinear, there exists at most one transformation t of G which carries $A \rightarrow A'$, $B \rightarrow B'$, $C \rightarrow C'$, $D \rightarrow D'$. The actual existence of such an element t cannot be proved until Axiom 6 is replaced by a slightly stronger one.

Let L be any line and P be any point. Let M and N be two lines concurrent with L neither of which passes through P. There is at most one element α of G such that α is in $(P \cap L)$ and α: $M \rightarrow N$.

AXIOM 6′: *The element α, described in the above theorem, exists. Axiom 6′ implies its dual and also implies Axiom 6. From Axiom 6′ it follows that there exists exactly one projective collineation which carries one set of four points, no three of which are collinear, into another set.*

The group G and the group of projective collineations of points and lines of G (i.e., its group of inner automorphisms) are isomorphic.

The Automorphisms of G.—The characterization of the projective collineation group we have just given is of a form which makes the study of its complete automorphism group extremely simple.[4] For this it is unnecessary to introduce coördinates. The study is carried out in two parts corresponding to the cases of odd and even dimensional spaces. The reason for this separation is that in a projective space of odd dimensionality there may exist collineations of period 2 without fixed points, and this slightly complicates matters. The special cases of projective spaces without collineations of period 2 (i.e., spaces whose coördinate fields are of characteristic 2) also have to be treated separately. We shall describe briefly the process for the two-dimensional case. The sequence of theorems is as follows:

A homology with a given center and axis is characterized by the fact that it permutes with every collineation of G which permutes with the harmonic homology with the given center and axis.

In any automorphic mapping of G, the image of a homology is a homology and conversely, every homology is the image of some homology.

The product of two homologies with different centers and different axes is not a homology.

If t_1 and t_2 are harmonic homologies with a common center A, their images in any automorphism of G are harmonic homologies with a common center or a common axis, and dually.

Two homologies with a common center are carried by an automorphism into two homologies with a common center or a common axis and dually.

Any automorphism of G carries an elation into an elation, and conversely, every elation is an image of an elation.

The only subgroups of G which consist of homologies and elations alone are points, lines and their subgroups.

Let P be any point. If an automorphism T makes $a \rightarrow a^T$ for each element a of G then P^T, the group of all images of elements of P, is a point or a line. The dual is also true.

If an automorphism T maps one point into a point it maps every point into a point and every line into a line; if T maps one point into a line it maps every point into some line and every line into some point.

MAIN THEOREM: *Let G be the projective collineation group of a projective plane π. With each automorphism of G we can associate a unique collineation or correlation (not necessarily projective) of π, and, conversely, with each collineation or correlation of π we can associate a unique automorphism of G, the mapping being an isomorphism.*

If one introduces a coördinate system into the geometry it can easily be seen that the complete collineation and correlation group is isomorphic to the group of all semilinear transformations in the coördinate field.

The case of n-dimensional geometry follows exactly the same lines once it is established that an automorphism maps a harmonic homology into a harmonic homology.

The Euclidean Subgroup of G (Two-Dimensional Case).—The identification of the Euclidean group as the subgroup of G which keeps a line and an absolute involution in it invariant has been established many times on geometrical grounds. From our point of view it is interesting to see if the Euclidean and orthogonal subgroups of G could be obtained directly from our group-theoretic definition of G. This in itself would be a difficult problem but for the work of Thomsen.[1] In essence, Thomsen gives a group theoretic characterization of the orthogonal group and what we are able to accomplish is to show that Thomsen's axioms are implied by ours when a proper specialization is made.

[*] This paper is a summary of a thesis presented at the University of Toronto in June, 1942, for a Ph.D. degree. The work was done under the supervision of Prof. G. deB. Robinson, for whose many aids and suggestions my thanks are extended.

[1] Thomsen, *Math. Zeit.*, **34,** 668–720; **37,** 561–566.

[2] Hjelmslev, J., "Einleitung in die allgemeine Kongruenzlehre," *Det Kgl. Danske Videnskabernes Selskab. Math. Phys. Meddelser*, **8** (11) (1929) and **10** (1) (1929).

[3] Cf. Robinson, *Foundations of Geometry*, Toronto, 1940.

[4] This problem was first treated (in an entirely different manner) by Van der Waerden and Schreier, *Abh. Math. Sem. Hamb.*, **6,** 302–322 (1928).

PROCEEDINGS
OF THE
NATIONAL ACADEMY OF SCIENCES

Volume 30 · October 15, 1944 Number 10

Copyright 1944 by the National Academy of Sciences

*VISION AND GUANINE PRODUCTION IN FISHES**

BY FRANCIS B. SUMNER

THE SCRIPPS INSTITUTION OF OCEANOGRAPHY

Communicated August 29, 1944

Zoölogists have long known that the purine body guanine, alone or in some combination, is responsible for the silvery appearance of areas of skin in most fishes. It is also a familiar fact that crystals of this substance are contained in special cells, variously called guanophores, leucophores or iridocytes, which respond, as do other chromatophores, to stimuli received through the eyes. The paling of a fish upon a white background results in part from the withdrawal from view of dispersed melanin granules within the melanophores, but also in part from the centrifugal movement of guanine particles within the guanophores. Reverse relations, with respect to both substances, are manifest upon a black background.

That the same stimuli which call forth these rapid adjustments may result, when long continued, in actual large-scale changes in the amount of pigment present has been shown by a number of workers.[1] For melanin, fairly exact relations have now been established in certain fishes between the amount of this pigment and the albedo of the background to which they have been subjected.

For guanine, no exact relations of this sort have hitherto been determined, though several workers have recorded the occurrence of an increase in the number of guanophores or in the quantity of visible guanine in animals kept in white containers.[2] In general, however, guanophores have received relatively little attention from students of chromatophore activity. This is doubtless largely due to their inconspicuousness in comparison with melanophores, likewise to their ordinarily being rendered invisible in histological preparations. The only worker known to me who professes to have undertaken the measurement of guanine in connection with color changes in animals (Meyer, op. cit.) reports a purely negative result, concluding that any possible experimental differences were swamped by the individual variability among the animals studied. This writer fails to state the method which she employed in making these measurements.

After the somewhat extended assays which have been made in this laboratory of the black pigment melanin, in relation to the visual environment, a comparable series of quantitative determinations of the white pigment guanine seemed called for.[3]

A local marine fish, *Girella nigricans*, was employed for this purpose. Two hundred and forty sexually immature[4] specimens were apportioned among 20 painted glass bowls.[5] Of the latter, there were four each of five shades, ranging as follows in respect to albedo, relative to white as a standard.

White	100.00
Pale gray	37.74
Medium gray	17.80
Dark gray	6.90
Black	1.42

Each set of four nearly identical bowls was kept in a separate cabinet, lighted by a 100-watt lamp overhead. Running sea-water was supplied to the bowls.

Of the original number of fishes, 217 were alive at the end of four months, when they were killed with chloroform vapor, dried with paper towels and weighed, commonly in lots of six or seven. These were then dipped into boiling distilled water for three-fourths of a minute and the "skins," including the scales and fins, were removed and dried for several hours in an oven at 100°C. These dried "skins" were also weighed to the nearest tenth gram.

The fishes from the first bowl in each cabinet, which were fewer in number and much larger than the others, were sacrificed for preliminary tests of method. The occupants of the other three bowls of each set were divided into two lots of six or seven fishes each. There were thus six lots from each cabinet (thirty lots in all) which were available for guanine assay. Of these six lots only five have been used.

Thus the fishes under consideration are designated as follows: "1," "2.1," "2.2," "3.1," "3.2," "4.1," "4.2". Of these the "1" lots were used for preliminary tests, while the "4.2" lots have not yet been used at all.

The matter is further complicated by the fact that the solid material of the original "2.1" series was divided into two equal portions ("2.1*a*" and "2.1*b*") and each of these was treated as an independent sample. The results from these are not quite comparable with those from the subsequent series, in that the extraction with acid from the former was less complete than from the latter.

For the methods which I have employed in the extraction and measurement of guanine I am indebted to Dr. George H. Hitchings of the Wellcome Research Laboratories, Tuckahoe, New York, who is reporting upon his own studies of some of the same material in another article in this issue of the PROCEEDINGS.[6] It is a pleasure to thank Dr. Hitchings for help and advice, kindly given at various stages of this undertaking, though I cannot hold him responsible for every minor step in my procedure. I must also acknowledge the important help of my colleague, Dr. Denis L. Fox, whom I have consulted frequently throughout these studies, as well as using some of the facilities of his laboratory. I must mention, too, the valuable technical assistance of Miss Ruth A. Larson during the earlier stages of these investigations.

The procedure here adopted was as follows:

1. The oven-dried skin was extracted with $N/1$ H_2SO_4, in the proportion of 10 cc. of the dilute acid per gram of the original wet weight of the fishes. Extraction here comprised grinding of the dampened material in a mortar, followed by four cycles[7] of prolonged shaking and centrifuging, in alternation, using a new fraction of the acid for each cycle.
2. Measured portions of this acid extract were heated for one-half hour in 100-cc. centrifuge tubes in a water bath.
3. These were made slightly alkaline with NaOH, using phenolphthalein as an indicator.
4. The contained purines were precipitated with copper sulfate, followed by sodium bisulfite, at boiling temperature.
5. The precipitate was brought down by centrifuge.
6. Precipitate washed twice with boiling water, followed each time by centrifuging.
7. Precipitate redissolved in boiling dilute HCl.
8. Copper precipitated by H_2S, passed through the heated solution.
9. Last precipitate removed from the guanine-HCl solution by filtration through a sintered glass filter.
10. Phenol reagent[8] (Folin-Denis-Ciocalteu) added to the filtrate, followed by Na_2CO_3.
11. Colorimetric determinations made of the resulting blue solution by means of a Bausch and Lomb Visual Spectrophotometer.

The degree of specificity to be expected from this method is discussed by Hitchings and Falco (pp. 294–297 of this issue). These workers have likewise added independent evidence, by the application of an altogether different method, that we are actually dealing here with guanine. This is important, since the procedure which I have employed is not sufficiently specific to distinguish between guanine and certain closely related compounds. It is gratifying to note that the values obtained by Hitchings and Falco, by the use of these two methods agree very closely; and it is further gratifying to note the reasonably close agreement between their figures and mine for the four samples which were assayed independently in the two laboratories by the copper precipitation method (Black 3.1, 3.2; White 3.1, 3.2). My figures should be compared with those in the middle column of Hitchings and Falco (0.285, etc.).

In making the spectrophotometer determinations, absorption values (logarithmic) were recorded by me at 550, 600 and 650 mμ. Colorimetric analysis of this sort obviously requires the previous plotting of graphs, based upon various dilutions of a standard guanine solution (Fig. 1). The values given in the tables, and in figures 2 and 3, are the average values obtained for these three wave-lengths. Five readings at each wave-length were comprised in each determination. It must be added that two to five subsamples of the acid extract were employed in the case of every lot of

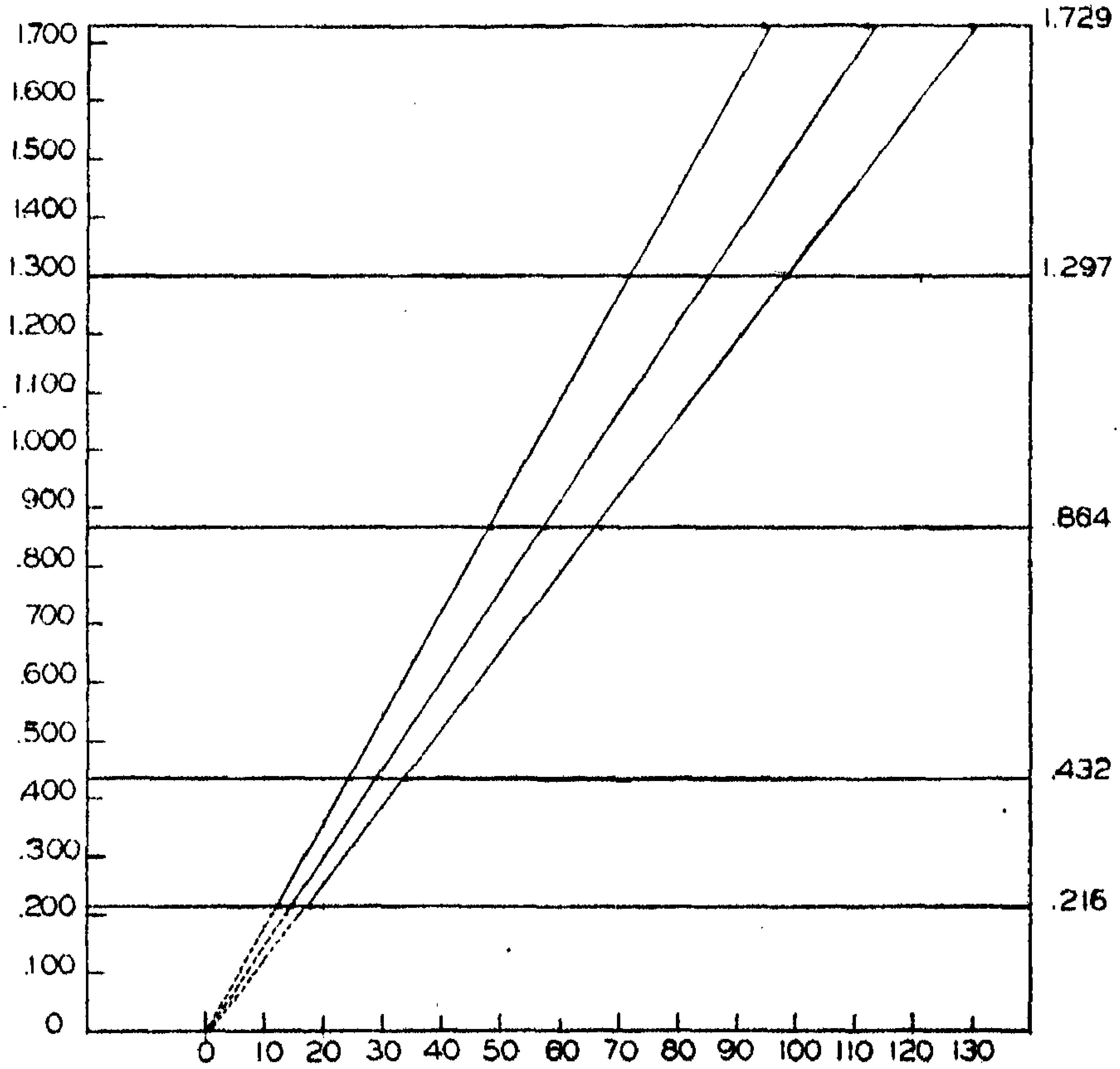


FIGURE 1

Readings with dilutions of standard guanine-HCl solutions. Abscissae = spectrophotometer readings (logarithms of absorption at 550, 600 and 650 mμ). Ordinates = mg. guanine per 100 cc. of solvent.

fishes, and that the means here presented are based upon all the values thus obtained, commencing with "2.1*a*," except for a few cases in which the treatment had been varied for special purposes.

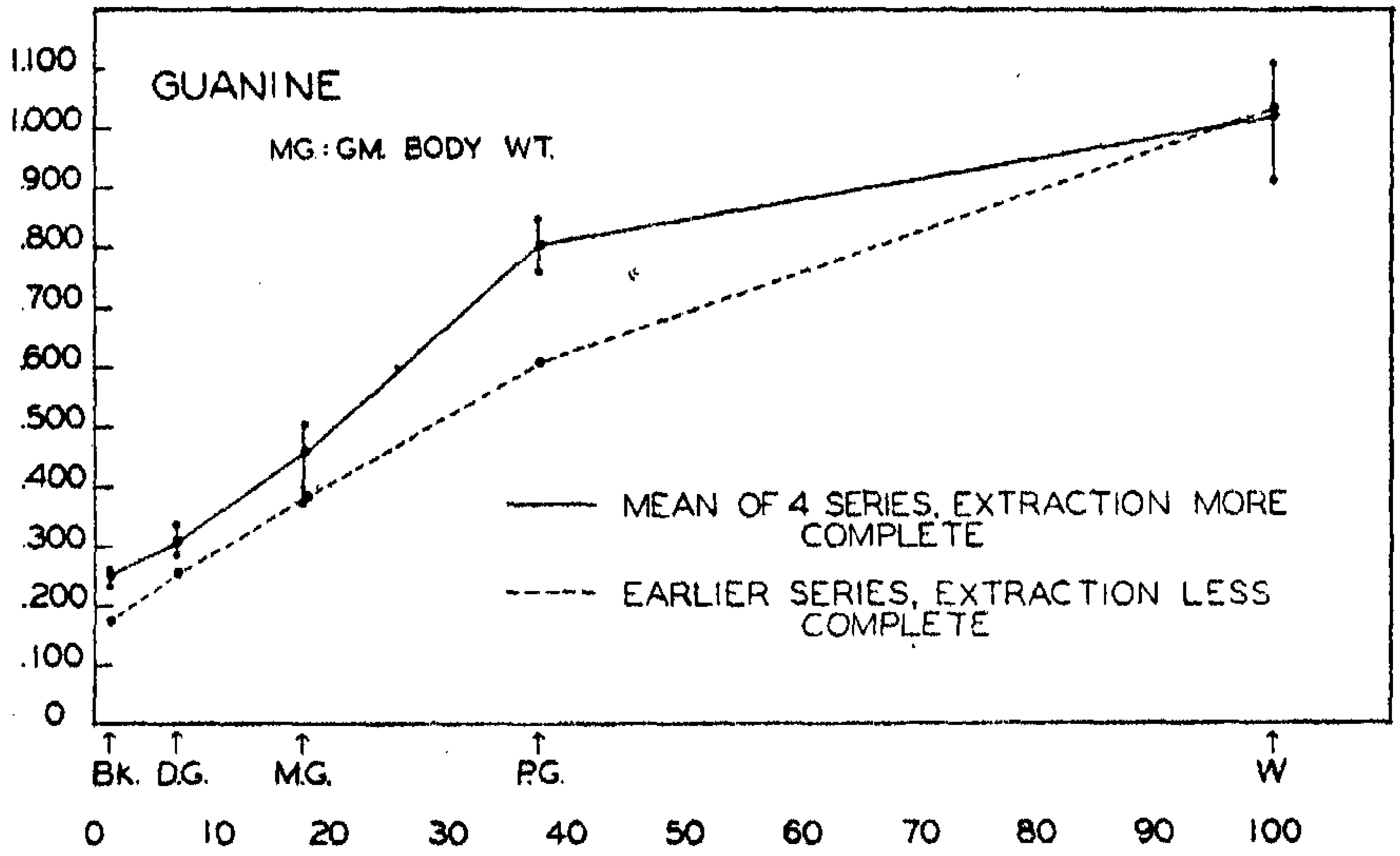


FIGURE 2

Guanine values for *Girella* series. Abscissae = albedos of the five shades of background. Ordinates = mg. guanine per g. of wet weight of fishes. Continuous line based on means of series "2.2," "3.1," "3.2," "4.1." Range of each series indicated by vertical bars. Broken line based upon means of "2.1*a*" and "2.1*b*."

TABLE 1

GUANINE, FIRST EXTRACTION

	BLACK		DARK GRAY		MEDIUM GRAY		PALE GRAY		WHITE	
	MG./G TOTAL WT.	MG./G DRY SKIN	MG./G TOTAL WT.	MG./G DRY SKIN	MG./G TOTAL WT.	MG./G DRY SKIN	MG./G TOTAL WT.	MG./G DRY SKIN	MG./G TOTAL WT.	MG./G DRY SKIN
2.1*a*	0.175		0.243		0.366		0.604		1.036	
2.1*b*	0.179		0.264		0.397		0.618		1.021	
Means	**0.177**		**0.253**		**0.381**		**0.611**		**1.028**	
2.2	0.231	3.61	0.286	4.27	0.371	5.80	0.756	10.65	0.972	13.89
3.1	0.260	3.71	0.293	4.58	0.483	6.90	0.766	11.27	1.084	15.27
3.2	0.239	3.62	0.300	4.48	0.504	6.00	0.851	11.35	0.913	13.63
4.1	0.256	3.77	0.333	5.05	0.467	6.97	0.851	12.33	1.110	16.32
Means	**0.246**	**3.68**	**0.303**	**4.59**	**0.456**	**6.42**	**0.806**	**11.40**	**1.020**	**14.78**

While occasional considerable differences were encountered between subsamples which should have yielded identical results, the mean deviation of such values from their average was 2.5 per cent (extremes: 0 and 6.6 per cent). Such differences are mainly small in comparison with the differences among the various experimental lots which had been subjected to different backgrounds.

The mean guanine values of the various lots of fishes herein discussed are presented in table 1. These values have been computed in two ways. In one column is presented the amount of guanine in the "*skin*" (as defined above) per gram of *total original wet weight* of the fishes. This may not seem to be a very logical value to employ, since the very considerable amount of guanine present in some other tissues of the body was not ordinarily known. In the other column we have the amount of guanine in the skin, in proportion to *the dry weight of the skin*. This last would doubtless be the preferable value to employ, except for the circumstance that the "skin," as here dealt with, was inevitably accompanied by a small but rather variable amount of adhering muscle. This would somewhat affect its weight, probably without correspondingly affecting its absolute guanine content. Thus the values in the first of these columns seem to furnish a fairer index of the changes in the concentration of this substance. Actually, the figures based upon the dry weight of the skin yield a graph very similar to those shown in figure 2.

No calculation of probabilities seems necessary in support of the relations shown by these graphs. Considering the four fully comparable series, "2.2," "3.1," "3.2," and "4.1," we find that the guanine content of fishes from the white bowls is 4.15 times that of those from the black ones, when ratios to total wet weight are considered; 4.02 when ratios to dry skin weight are considered.

In view of the readily visible guanine deposits in the peritoneum of these fishes, it is not surprising that considerable of this substance may be extracted from their "bodies" after removal of the "skins." Only two lots were tested for this purpose, but in these the total amount contained in the "bodies" was absolutely greater than that in the "skins" (1.06 and 1.19:1). Since the latter constituted only about 23 per cent[9] of the dry weight of the fishes, the concentration of guanine was obviously much greater in the skins.

It was early found that even after four "cycles" of extraction, as above described, the fish material continued to yield some substance which gave distinct traces of the blue reaction with phenol reagent and sodium carbonate. The material from the "4.1" series was subjected to these additional extractions, these being carried through four cycles, as had been done originally. The results are shown in table 2.

The mean absolute values of these "second extractions" thus do not differ significantly for the different lots. Indeed, their magnitude appears to bear no relation to the source of the chromogen, a circumstance which is in striking contrast to the figures for the "first extractions," which range from 0.256 to 1.110, as we pass from black to white. This fact suggests that we are dealing with a different chromogen in the earlier and later extractions, and that in these last the amount is not influenced by optical conditions.[10]

I may add that a third extraction (i.e., series of extractions) yielded 0.009 and 0.008, respectively, for "black" and "white" lots (the other lots were not tested). These last figures are of about the same magnitude as was obtained when the H_2SO_4 alone, of the same dilution, was employed, without any known contamination by organic matter. It would seem that all of these guanine measurements are subject to a small constant error, which perhaps could be determined more precisely and applied to the results presented in the tables.

Several earlier writers have stressed the reciprocating relations between guanophores and melanophores, resulting from changes in the background, but without attempting to measure either of the pigments concerned. Thus Kuntz (1917) remarks that "complete adaptation to a white background involves a rearrangement and probably an increase in the number of the guanophores and a material reduction in the number of melanophores in the layer just beneath the epidermis" (op. cit.).

TABLE 2

GUANINE, SECOND EXTRACTIONS

	FIRST EXTRACTIONS	SECOND EXTRACTIONS 1	2	MEAN	% OF FIRST EXTRACTION
Black	0.256	0.021	0.017	0.019	7.4
Dark gray	0.333	0.024	0.020	0.022	6.6
Medium gray	0.467	0.024	0.024	0.024	5.1
Pale gray	0.851	0.022	0.019	0.021	2.5
White	1.110	0.025	0.018	0.021	1.9

Murisier (1920–1921) notes that melanin and guanine seem to be substances which may substitute for one another, the amount of one which is formed being inverse to that of the other. Like melanin, it is formed under the influence of retinal stimuli. Murisier suggests that both guanophores and melanophores form an accessory excretory system, but one which retains its products instead of discharging them.

Millot (1923) sets forth these same ideas at greater length: "The three pigments (melanin, guanine, lipochrome) have almost the same physiological value and belong to the category of waste materials" (trans.). Although guanine is an excretory product, its abundance in the tissues is altogether independent, he asserts, of the food consumed or of the activity or integrity of the renal apparatus.

"Experiment reveals in a number of cases the existence of an equilibrium between the melanin and the guanine such that any diminution of the quantity of one pigment brings about a compensatory augmentation of the other."

While this reciprocal relation between guanine and melanin has been affirmed by several previous writers, it has never, so far as I know, been supported by quantitative data. Although Millot discusses a method for isolating guanine, he gives no figures to support his claims that the amount of this substance may be influenced by visual stimuli.

In figure 3 of the present paper I have placed in comparison the curve for guanine production in *Girella nigricans* with that for melanin production in this same species, which was obtained under almost identical experimental conditions.[11] The approximate character of the values upon which both of these curves are based has been freely admitted.

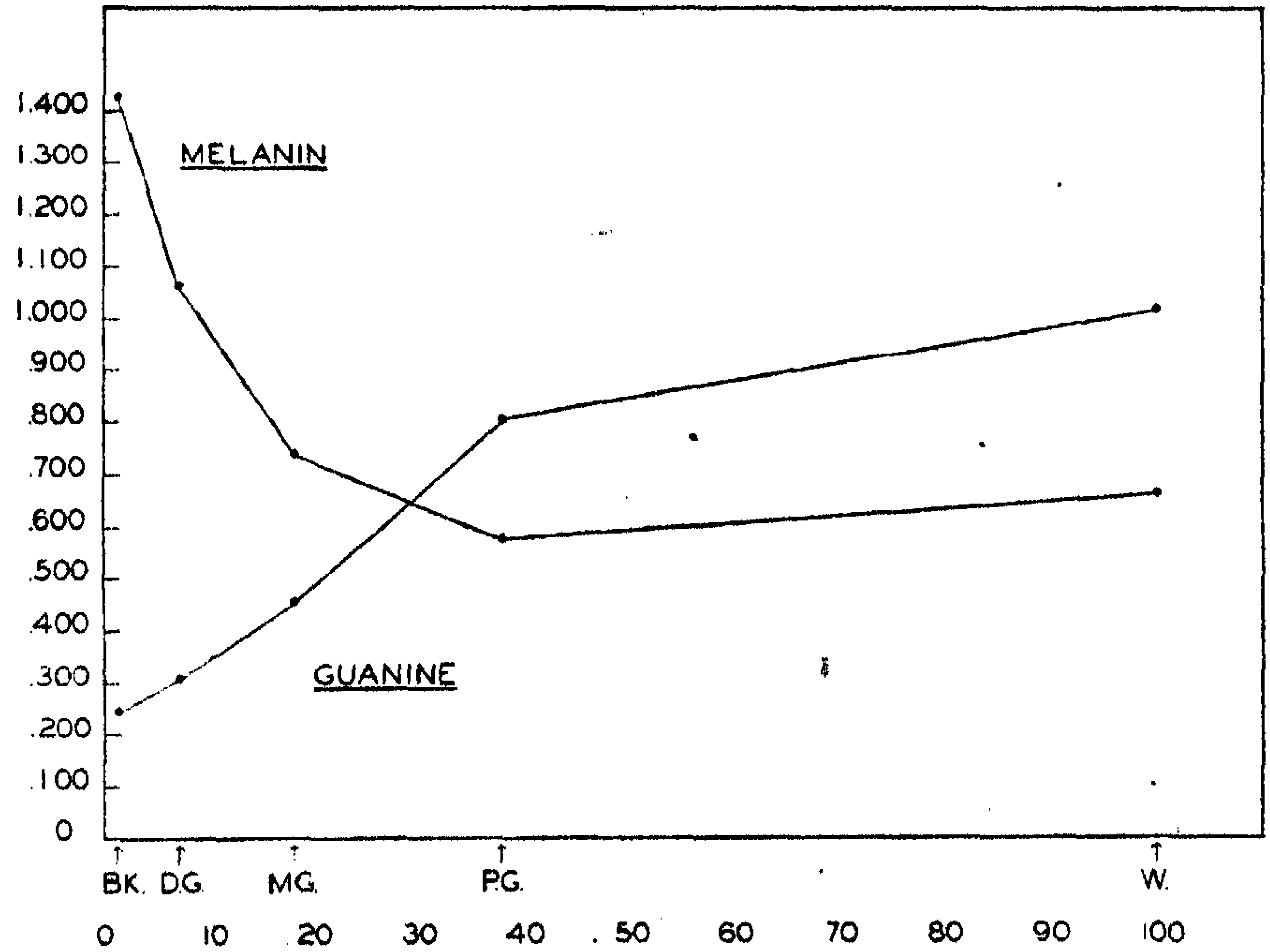


FIGURE 3

Comparison of curves for guanine and melanin in *Girella*. Guanine curve = continuous line of figure 2. Melanin curve is taken from Sumner, 1943.

While the melanin values, from black to pale gray, vary inversely as the logarithm of the albedo, no such arrangement holds for the guanine values. However, the two curves agree in displaying a marked change in their trend beyond pale gray, that for melanin ceasing to fall (indeed, rising slightly) at this point, that for guanine undergoing a marked reduction in the rate of increase. It is, of course, unfortunate that an intermediate albedo between "pale gray" and white was not intercalated in this series, but the relative length of this interval was not realized by eye at the time of mixing the paints.

The outstanding fact revealed by these experiments is that the course of protein metabolism, and the production of two important metabolites, melanin and guanine, are strongly influenced by the visual field of the fish, and influenced in such a way as to make for the animal's concealment.

These fundamental organic processes are thus caused to change their course in the interest of superficial appearance, and this with no apparent effect upon the growth of the fishes. It is true that the strict accuracy of this last statement should be tested more carefully than was done in the present experiments, in which the original mode of selection of the animals, at the time they were distributed to the various bowls, appears to have led to minor size differences which are irrelevant to the question at hand.

That this mechanism for concealment may be of life-saving value has been shown by considerable experimental evidence.[12]

If these two pigments (three, if we follow Millot) actually function as excretory products, it is surely surprising that the necessary excretory balance should be maintained through reciprocal adjustments in the production of substances so widely unrelated chemically. Plainly, there are physiological (shall we not say psychobiological?) problems here which call for extensive future investigation.

---

[*] Contributions from the Scripps Institution of Oceanography, New Series, no. 235,

[1] A summary of observations in this field was published by Sumner, *Biol. Rev.*, **15**, 351–375 (1940). Some later results have been reported by Dawes, *Jour. Exper. Biol.* **18**, 26–49 (1941); by Sumner and Doudoroff, *Biol. Bull.*, **84**, 187–194 (1943); and by Sumner, *Biol. Bull.*, **84**, 195–205 (1943).

[2] Kuntz, *Bull. U. S. Bur. Fisheries*, **35**, 1–29 (1915–1916); Murisier, *Revue Suisse de Zoologie*, **28**, 45–97, 149–195, 243–299 (1920–1921); Millot, *Bull. Biol. de la France et de la Belge*, **57**, 261–363 (1923); Hewer, *Phil. Trans. Roy. Soc. London*, **B215**, 177–200 (1927); Meyer, *Zool. Jahrb.* (*Abt. f. allgem. Zool.*), **49**, 231–270 (1931); Sumner and Wells, *Jour. Exper. Zool.*, **64**, 377–403 (1933); Odiorne, *Proc. Nat. Acad. Sci.*, **19**, 750–754 (1933).

[3] I think that guanine may here be legitimately rated as a *pigment*, inasmuch as it actually functions as a pigment, i.e., influences the *color* (in the broad sense) of the animal.

[4] This point is important to note, since it makes certain that no sexual differences in pigmentation can have complicated the picture.

[5] This is the same set of bowls that was employed in the writer's melanin experiments with the same species (*Biol. Bull.*, **84**, 197 (1943)). They were of clear, colorless glass, painted on the outside, and sand-blasted within to eliminate gleams of reflected light.

[6] Hitchings and Falco, *Proc. Nat. Acad. Sci.*, **30**, 294–297 (1944). See also Hitchings, *Jour. Biol. Chem.*, **139**, 843–854 (1941); Hitchings and Fiske, **140**, 491–499 (1941).

[7] Only one cycle in the case of "2.1*a*" and "2.1*b*."

[8] Folin, *Laboratory Manual of Biochemistry*, D. Appleton-Century, 1934; Folin and Ciocalteu, *Jour. Biol. Chem.*, **73**, 629 (1927).

[9] Based upon only one lot of 3 fishes.

[10] Of course, a certain amount of this second chromogen may be supposed to occur in the earlier extractions.

[11] The latter from Sumner, *Biol. Bull.*, **84**, 195–205 (1943).

[12] Sumner, Amer. Nat., **69**, 245–266 (1935); *Proc. Nat. Acad. Sci.*, **20**, 559–564 (1934); **21**, 345–353 (1935).

THE IDENTIFICATION OF GUANINE IN EXTRACTS OF GIRELLA NIGRICANS. THE SPECIFICITY OF GUANASE[*]

By George H. Hitchings and Elvira A. Falco

The Wellcome Research Laboratories, Tuckahoe, N. Y.

Communicated August 29, 1944

Sumner in the preceding article (pp. 285–294) has shown that the apparent concentration of guanine in the skin of *Girella nigricans* varies with the albedo of the background against which the fish are kept.

Guanine was identified as a constituent of fish scales about eighty years ago[1,2] and commercial "pearl essence" from this source is believed to consist essentially of free guanine.[3] The existence of guanine in fish scales may therefore be accepted, though with some reservations, since the chemical identification was carried out before isoguanine (2-hydroxy-6-aminopurine) which has properties nearly identical to those of guanine,[4] became available. Within recent years a new series of pigments closely allied to the purines has been studied in detail by Wieland and associates.[5,6] These include isoguanine ("guanopterin")[7] and a number of derivatives of the pteridine (azinepurine) heterocycle. These pigments were isolated first from butterfly wings, but at least one member now is known to exist in fish skin.[8] Therefore, though the identification of guanine as a major constituent of fish skin and scales may be accepted, the absence of similar substances is by no means certain.

The procedure adopted by Dr. Sumner for the determination of guanine, involves a preliminary separation of the material by a precipitation with copper and bisulfite, and finally a colorimetric estimation with phenol reagent. The copper-bisulfite precipitation limits the number of substances which may be present to those of the purine group and substances closely allied chemically[9] but the reaction with the phenol reagent is relatively unspecific.[10] Isoguanine, for example, and some of the pteridines might be expected to be determined as guanine by these procedures. Therefore a quantitative chemical or biochemical procedure was sought which would increase the specificity of the method. Ideally such a procedure, applied to small amounts of material, would distinguish between guanine and all other compounds known to behave like guanine with respect to copper precipitation and the phenol reagent. This would enable one to determine what proportion of the material estimated as guanine was authentic guanine. Above all one could then determine whether the observed changes in concentration of the substance due to environmental conditions were due chiefly to changes in the concentration of guanine. To a considerable extent a procedure including the use of the enzyme guanase[11] was found to give the desired degree of specificity. This enzyme has been used by Schmidt[12] for the determination of guanine in tissue extracts, but its action on substances closely related to guanine had not been studied previously. Consequently a study of the specificity of guanase action was undertaken.

Guanase was prepared from acetone-dried rabbit liver powder by extraction with water. The extract was purified by treatment with dialyzed iron and finally the enzyme was adsorbed on alumina cream,[13] and eluated with a 0.2 *M* disodium phosphate solution. An amount of the substance to be tested equivalent to 0.1 mg. of amino nitrogen was incubated 18 hours at 37° with guanase in 0.1 *M* phosphate buffer at pH 7.8. After the addition of an excess of sodium tetraborate solution, the ammonia which had been liberated was distilled into standard 0.02 *N* acid and the excess acid was titrated. The action of guanase on twelve purines and azine-purines closely related to guanine was studied. Within the limits of experimental error, the enzyme showed no liberation of ammonia with the following substances: adenine, 7-methylguanine, 1, 7-dimethylguanine, 2-amino-6,8-dihydroxypurine ("8-hydroxyguanine"), 2-hydroxy-6-aminopurine ("isoguanine"), 2-amino-6-hydroxyazinepurine, xanthopterin, isoxanthopterin, xanthopterin carboxylic acid and isoxantopterin carboxylic acid. In 10 experiments with guanine an average of 98 per cent of the theoretical amount of ammonia was liberated (range 93–106 per cent). Guanase showed a definite action on 1-methylguanine. In three experiments carried out as above an average of 81.0 per cent (maximum 84.2 per cent) of the theoretical nitrogen was liberated. The action of guanase was demonstrably slower on 1-methylguanine than on guanine itself. With a five-hour incubation period and a somewhat feebler guanase solution, 69 per cent of the theoretical ammonia was liberated from guanine, while 8.5 per cent was being liberated from 1-methylguanine.

Action of Guanase on the Purine of Extracts of *Girella Nigricans*

	GUANINE FOUND, MG. GUANINE PER G. FISH		
	DIRECT	AFTER COPPER PRECIPITATION	
SAMPLE*	COLORIMETRIC	COLORIMETRIC	GUANASE
Black 3.1	0.386	0.285	0.301
3.2	0.347	0.250	0.245
White 3.1	1.23	1.11	1.08
3.2	1.04	0.91	0.85

* Designations of Dr. Sumner (pp. 285–294).

The action of guanase on the purine fraction of extracts of *Girella nigricans* then was studied. For this purpose fish were chosen to represent wide differences in apparent guanine content. The results of these experiments are given in table 1. Each extract contained about the same amount of material which reacted with the phenol reagent but was not precipitated by copper and bisulfite and thus was non-purine in character. In each instance the concentration of guanine found by the guanase method was identical with that found colorimetrically within the limits of accuracy of the enzymatic method.

Whereas guanase failed to show a complete specificity in its action, by its use it was shown that isoguanine and a number of purines and pteridines could not be present in the extracts. The failure of guanase to distinguish in all or none fashion between 1-methylguanine and guanine,[14] leaves open perhaps some possibility that the former substance might be present in the extracts, though 1-methylguanine has never been isolated from natural sources. To a considerable extent, this possibility was eliminated by a determination of the substance as the picrate.[15] A solution of the fish purine containing 0.92 mg. total nitrogen, and 0.81 mg. guanine nitrogen as determined colorimetrically was precipitated at a volume of 6 ml. by half saturation with sodium picrate. Titration of the precipitate gave an apparent guanine content of 0.85 mg. of nitrogen. When 1.00 mg. of guanine nitrogen was treated in the same way, 1.02 mg. were found by titration. On the other hand 1.00 mg. of nitrogen as 1-methylguanine gave a recovery of only 0.71 mg. due to the greater solubility of the picrate of this compound.

The results leave little reason to doubt that substantially all the apparent guanine of the extract is authentic guanine, and that the amount of guanine in the fish skin changes with changes in the environmental conditions under which the fish are kept.

---

* From The Wellcome Research Laboratories, Tuckahoe, New York.

[1] Barreswil, *Liebig's Ann.*, **122**, 128 (1862).

[2] Bethe, A., *Hoppe Seyl. Z.*, **20**, 474–477 (1895).

[3] Tressler, D. K., *Marine Products of Commerce*, Chemical Catalog Co., New York, 1923, pp. 161–166.

[4] Fischer, E., *Ber. dtsch. chem. Ges.*, **30**, 2245 (1897).

[5] Wieland, H., and Schöpf, C., *Ibid.*, **58**, 2178–2183 (1925).

[6] Schöpf, C., and Becker, E., *Liebig's Ann.*, **524**, 49–123 (1936).

[7] Purrmann, R., *Ibid.*, **544**, 182–190 (1940).

[8] Hüttel, R., and Sprengling, G., *Ibid.*, **554**, 69–82 (1943).

[9] Hitchings, G. H., and Fiske, C. H., *Jour. Biol. Chem.*, **140**, 491–499 (1941).

[10] Hitchings, G. H., *Ibid.*, **139**, 843–853 (1941).

[11] Schmidt, G., *Hoppe Seyl. Z.*, **208**, 185–224 (1932).

[12] Schmidt, G., and Engel, E., *Ibid.*, **208**, 225–236 (1932).

[13] Marshall, J., and Welker, W. H., *Jour. Amer. Chem. Soc.*, **35**, 820–822 (1913).

[14] It is interesting to note in this connection that in another biological activity, the

germination of spores of *Phycomyces blakesleanus*, 1-methylguanine approaches in activity guanine itself where other methyl guanines and closely related compounds have activities of a lower order of magnitude. (Unpublished experiments with Wm. J. Robbins.) Since methylation in the 1-position would result in fixation of the lactam tautomer, it is perhaps permissible to speculate that with guanase and with *Phycomyces* it is the lactam form of guanine which is active.

[15] Hitchings, G. H., and Fiske, C. H., *Jour. Biol. Chem.*, **141**, 827–835 (1941).

ON SPONTANEOUS MUTATION

By Richard Goldschmidt

Department of Zoölogy, University of California

Communicated August 3, 1944

The occurrence of spontaneous mutation in a stock called *px bl* was analyzed. Some of the facts have been reported previously,[1] but there were some errors of interpretation in the preliminary report. The *px bl* line is derived from standard plexus by mutation at the *bs* locus together with an appearance of modifiers for plexation and blistering. In an inbred and closely watched brother-sister culture an "upheaval" (formerly called mass-mutation) occurred, in which mutants at the silver arc and rudimentary loci occurred simultaneously, together with return mutation of *px* and *bs* to wild. A similar upheaval was found twice later in mass culture, the same loci being involved but in part with different alleles. All these mutants and return mutants appeared individually many times later in controlled broods of derivatives of *px bl*, or in in-and-out crosses of *px bl* derivatives. These mutants include a long series of characteristic alleles at the *svr* and *a* loci, repeated mutation at the *bs* and *bb* loci, and further secondary return mutation from + to *px* and of the *svr* and *a* alleles to wild type.

In a significant number of cases, two or more mutants or return mutants, including mutation from one allele to another, occurred simultaneously. This applies especially to the *svr* and *a* alleles. In such cases abnormal sex ratios were frequent. With the exception of these occurrences, the mutation rate was very low in *px bl* and its derivatives. But after in-or-out-crossing this rate rose considerably and also an accumulation of mutants was observed within a few generations after the first mutant had been found. This includes not only mutants and return mutants at the loci already mentioned, but also the rather frequent occurrence of others at the Lobe, dachs, Notch, ebony loci. It is remarkable that in a somewhat small, tentative set of x-ray experiments an unusually high rate of mutation was observed, and among the mutants were alleles of *svr* and *a* as well as at the *bs*, *N* and *d* loci; of course, all lines were checked before the experiments for the absence of such mutants.

The *px bl* line as well as the lines of derived mutant stocks were thoroughly analyzed genetically and cytologically. Aside from the presence of typical modifiers for the expression of the main loci, a series of non-reciprocal, very small translocations from the *X*- into the autosomes was found, producing male lethal classes but otherwise viable. They did not contain known loci but could be localized approximately. The data show the insertion of small sections from near the left end of the *X*-chromosome into the second near arc and into the third near its right end. In addition, the genetic data require transpositions of the same *X*-chromosome loci, one into the right half between *v-wy*, the other to somewhere near the right end. Some but not all of these extremely small rearrangements could be found in the salivaries. The original *px bl* stock already contained these features in different combinations. There is a strong suspicion that a relation exists between the presence of these features and mutation at the *svr* and *a* loci. Strangely enough, some of the rearrangements were also found in an *svr* allele derived independently (in heat experiments) from Florida stock. A statistical check of the salivary chromosomes revealed the absence or the great rarity of conspicuous rearrangements for *px bl* and the chief *svr* allele derived from *px bl* in the first upheaval. But a very large number of 1–3 band deficiencies and insertions could be found, some frequently, some seldom; they include the insertions from *X*- into autosomes already mentioned. In the series of *svr* alleles the locus (left of the *Bld* break) attributed to them through deficiency tests is always normal. In some of the alleles, however, a disturbance is typically found six bands to the right. In one allele it is a clear two-band inversion. In others the best possible interpretation of the abnormality is a one-band inversion. In the other long series of mutants at (or near) the *a* locus, only one shows a one-band deficiency. The other mutants derived from the line had either normal salivaries or were small deficiencies for known loci (Notch and black deficiencies). There is no reason to assume that the visible rearrangements were more frequent than the so-called point mutations.

The facts in this closely followed case of spontaneous mutation, together with a detailed study of the lines involved, tend to show that mutation is not a haphazard phenomenon but the consequence of preëxisting conditions in the chromosomes which act upon other points in the same or other chromosomes. Such facts—with which many data from existing literature may be considered as falling in line—suggest that mutation is not a chemical change within a gene-molecule, but a mechanical happening, probably a break under stress, produced by the synaptic properties of small rearrangements, including such breaks and rearrangements as are below the level of a single salivary band; i.e., so-called point mutants. The latter can now be envisaged after the intimate structure of one band has been elucidated by Kodani.[2] A full report is ready for press.

[1] Spontaneous chromatin arrangements and the theory of the gene, *Proc. Nat. Acad. Sci.*, **23**, 621–623 (1932).

[2] Kodani, M., "The Structure of Salivary Gland Chromosomes of *Drosophila melanogaster*," *Jour. Hered.*, **33**, 114–133 (1942). Goldschmidt, R., and Kodani, M., "The Structure of the Salivary Gland Chromosomes and Its Meaning," *Am. Nat.*, **76**, 529–551 (1942).

ONE-BAND INVERSION

By Richard Goldschmidt and Aloha Hannah

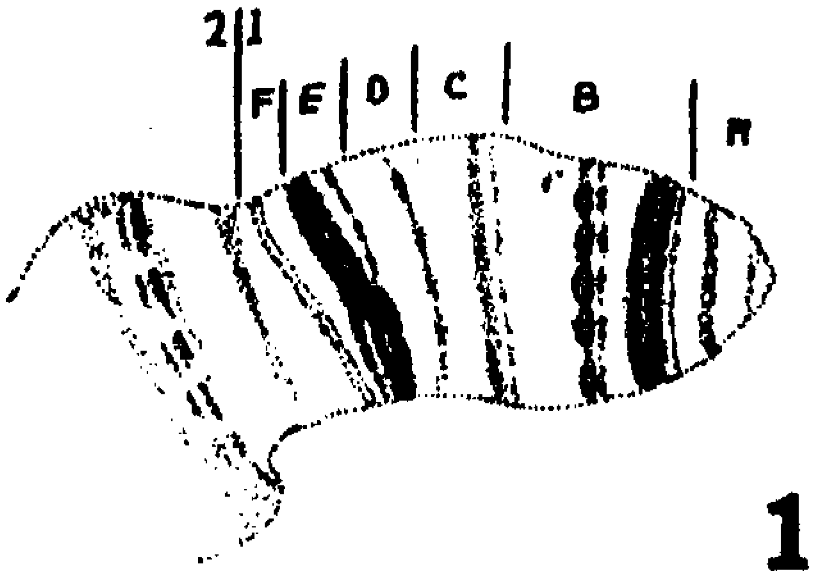
Department of Zoölogy, University of California

Communicated August 3, 1944

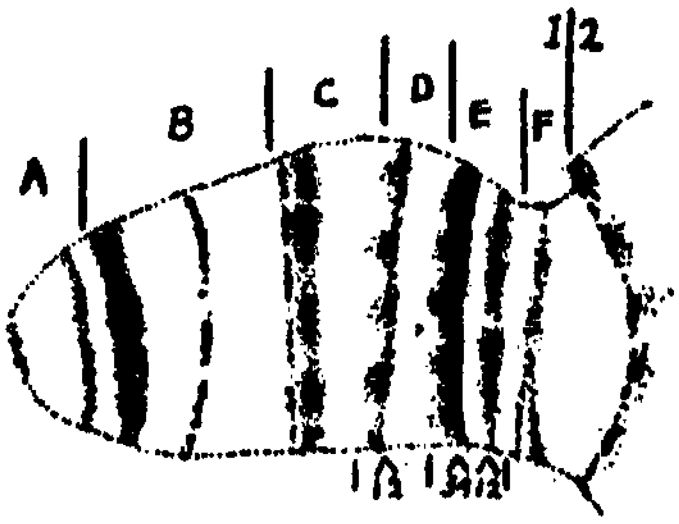
In a recent paper by Oliver and Green[1] the following remarks are found: "In some of the females heterozygous for lz^8, some but not all of the cells of the salivary glands show a torsion of one heavy band which is suggestive of a small inversion or possibly a duplication." The feature was found only in the presence of the allele lz^8. It happens that we have been studying in detail what probably is a parallel case. The mutants in question are a series of alleles at the silver locus (1–0.1) of *Drosophila melanogaster* which were studied closely in connection with work on spontaneous mutation.

On the basis of deficiency tests the silver locus is expected to be located to the left of 1*C* in the first chromosome. In all alleles, including those produced by x-rays, this region is perfectly normal. But in some alleles the region farther to the right, between the *Bld* break and the bulb, shows typical disturbances. The only one of these which can be described with certainty is found in the allele $svr^{pot\ s}$. Here a 2- (4-) band inversion of 1*E*1-4 is always found. These bands are described in Bridges' standard map (1938) as two double bands, the one to the left being thinner, that to the right being rather thick. We think that there exist actually only two bands, the double nature of which is not always visible, being based upon a partial separation of the two groups of perultimate chromomeres of which the bands are constituted (see Kodani, 1942; Goldschmidt and Kodani, 1943).[2] In the heterozygote this inversion looks as pictured in figure 1. One cannot always be sure that the interpretation as a minute inversion is correct. But in the homozygote the order of the thin and thick bands 1–2, 3–4 is clearly reversed, as shown in figure 2. Therefore we feel confident that the interpretation is correct.

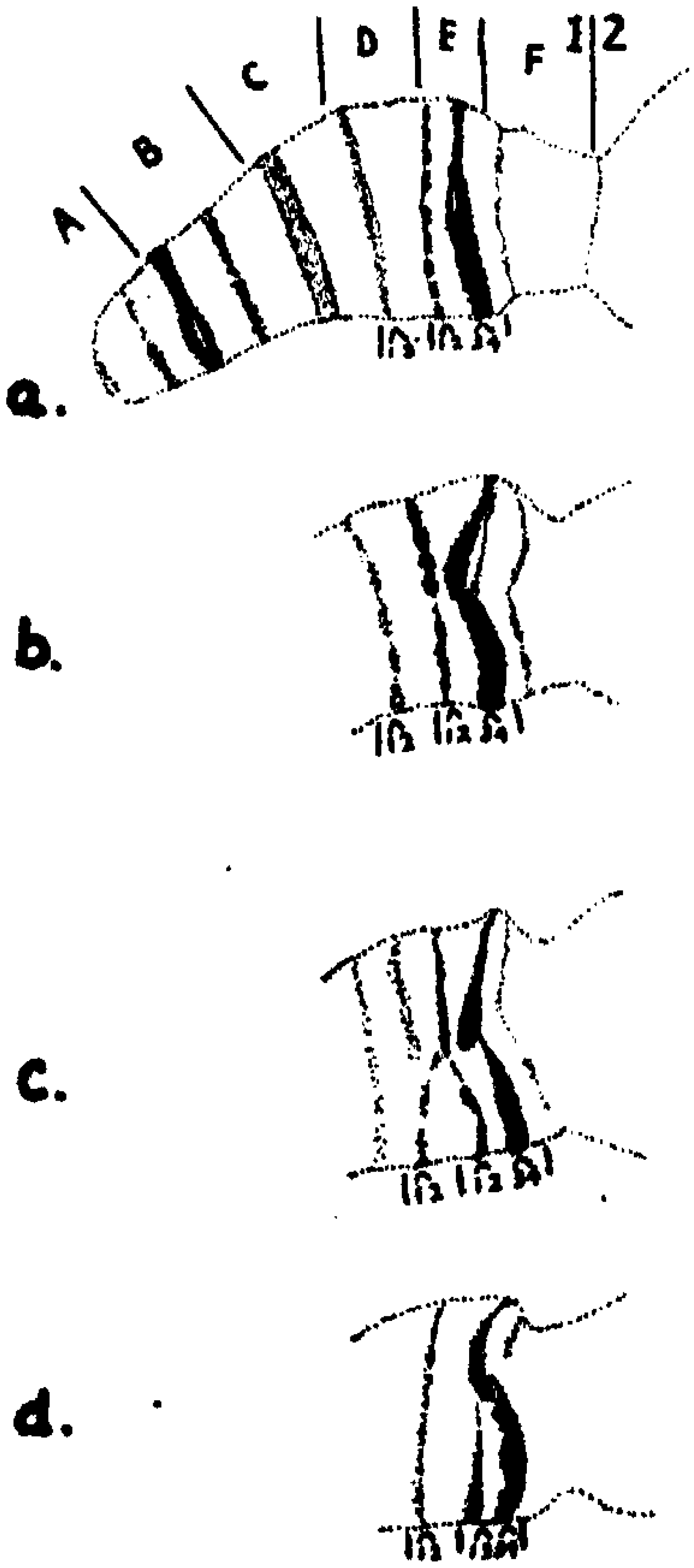
In some of the other alleles, namely, svr^{pot}, $svr^{pot\ sq}$, $svr^{pot\ duh}$, an abnormality of the same bands is found in a majority of *X*-chromosomes (this is never true of the controls), a structure which offers great difficulties of interpretation. The order of the bands 1–2, 3–4 is normal. But the thick



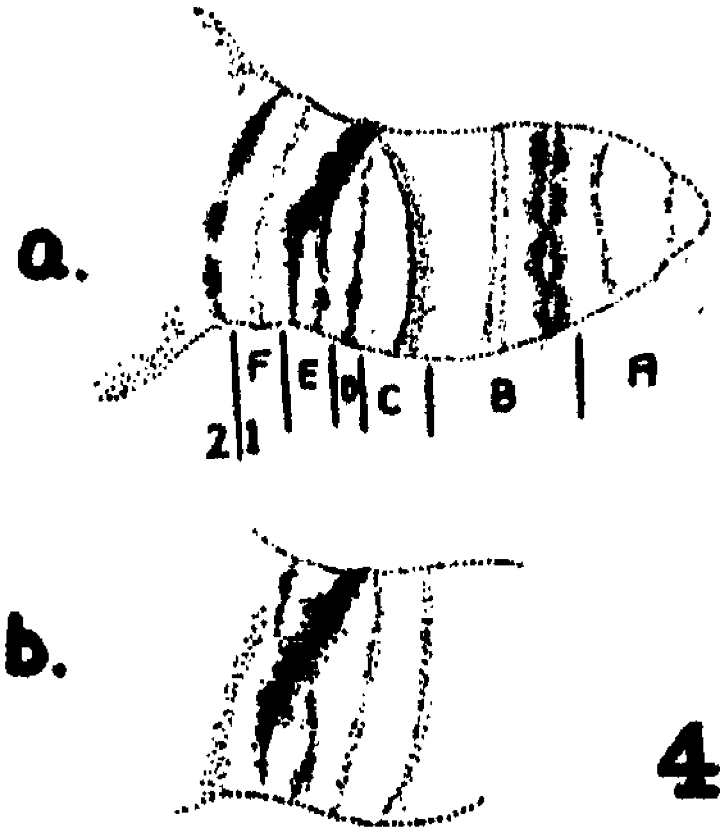
1



2



3



4

band 3–4 is distorted in a way which is difficult to describe. When focusing high and low the band seems normal, but in between a kind of cross or figure-eight appears, an apparent crossing of 1, 2 with 3, 4. This seems to be the result of two features; first, the arrangement of the chromatin on the surface of the disc-like band (like a ring), and secondly, a dislocation of the two halves of this ring belonging to the two homologous chromosomes, by a twist in opposite directions. This means that viewed from above the upper rim of the disc in one chromosome is shifted to the right, and that of the other chromosome to the left, so that a kind of figure-8 becomes visible. This is pictured in figures 3 and 4, where the same chromosome is drawn at different foci. Our interpretation of these typical aspects, which can hardly be drawn as they appear in focusing, is that the distortion within the disc represents a one-band inversion. This requires the assumption that the left and right surfaces of a single disc are serially different. The new discoveries on the structure of the individual disc (see Kodani, and Goldschmidt and Kodani[2]) make such an interpretation possible, since they revealed the presence of a coil of genonema with an attached series of perultimate chromomeres in a disc (each one double). An inversion of one disc would change the direction of the coil, creating a tension in the state of synapsis. (Consult figure 6 in Kodani's paper.) Homozygous chromosomes were always normal.

In $svr^{poi\ h}$ a different structure is found, again for the same bands. We did not succeed in finding a satisfactory interpretation. The only certain fact is that these bands are not normal. Drawings will be given in a forthcoming full report.

If our interpretation is correct—and we have no doubt that it is—the group of mutants which turn out to be small rearrangements will probably be considerably enlarged. Simultaneously the senior author's interpretation of the experimental work on the structure of the salivary chromosomes[2] will receive important support.

---

[1] Oliver, C. P., and Green, M. M., "Heterosis in Compounds of Alleles in *Drosophila melanogaster*," *Genet.*, **28**, 331–387 (1944).

[2] Kodani, M., "The Structure of Salivary Gland Chromosomes of *Drosophila melanogaster*," *Hered.*, **33**, 115–133 (1942). Goldschmidt, R., and Kodani, M. The Structure of the Salivary Gland Chromosomes and Its Meaning, *Amer. Natur.*, **76**, 529–551 (1942).

CHROMOSOME ASSOCIATIONS IN TETRAPLOID HYBRIDS BETWEEN PRIME TYPE 1 AND PRIME TYPE 2 IN DATURA STRAMONIUM

BY A. D. BERGNER*

UNITED STATES DEPARTMENT OF AGRICULTURE, SPECIAL GUAYULE RESEARCH PROJECT, SALINAS, CALIF.

Communicated September 8, 1944

Prime types among diploid plants of *Datura stramonium* have been discussed in several previous publications.[1, 2, 3] They are the result of interchanges between non-homologous chromosomes. No studies have been made of chromosomal configurations in the F_1 hybrids between tetraploid prime types (*PT*'s) in *Datura* with the exception of *PT* 1 × *PT* 2.

In diploid races which are homozygous for *PT*'s, the 24 chromosomes are associated as 12 bivalents at the meiotic first metaphase (*M* I) in pollen-mother-cells (*PMC*). If homozygous *PT* 1 (standard Line 1) is crossed with any other homozygous *PT*, the chromosomes of the hybrid show a configuration of four or more chromosomes plus 10 or fewer bivalents, depending upon which *PT* is crossed with *PT* 1. In tetraploid races which are homozygous for *PT*'s, the 48 chromosomes are associated as 12 quadrivalents, or the equivalent, at *M* I since there are four of each kind instead of the two in diploids. In a tetraploid hybrid between *PT* 1 and any other *PT* many different types of configuration are possible.

Diploid Prime Type 2.—*PT* 2 has the chromosomes **1·18**[4] and **2·17** instead of the 1·2 and 17·18 chromosomes of *PT* 1. The hybrid between *PT* 1 and *PT* 2 shows a $\odot$ 4 + 10 bivalents (*BV*) at *M* I. The chromosomes in the $\odot$[5] are as follows: (1·2-**2·17**-17·18-**18·1**).[6] Occasionally instead of the $\odot$ 4, which is illustrated in figure *A*, these four chromosomes form a chain (*ch*) of -**18·1**-1·2-**2·17**-17·18 or more rarely -2·1-**1·18**-18·17-**17·2**. They may also occur as the two pairs -**2·17**-17·18 + **18·1**-1·2. The $\odot$ 4 and *ch* 4 are usually so oriented on the spindle that disjunction is alternate.

Under favorable conditions these arrangements can be determined because of size differences among these four chromosomes. The *L*-sized 1·2 chromosome is twice as long as the *m*-sized 17·18 chromosome; **1·18** is slightly longer than 17·18 and **2·17** is slightly shorter than 1·2. These differences in length were first determined by using the location of the extra chromosome in certain $2n + 1$ types heterozygous for *PT* 2 at *M* I in *PMC*[7] and later checked by measuring these chromosomes in pollen grains.[8] The **1·18** and 17·18 chromosomes probably have the same insertion region; correspondingly the 1·2 and **2·17** chromosomes have the same insertion region. This conclusion is based on a comparison of total length and length of each arm of these four chromosomes in *PMC* and in pollen grains. The ·18 arm of the **1·18** chromosome, therefore, is the same as the ·18 arm of the 17·18 chromosome and the ·2 arm of the **2·17** chromosome is the same as the ·2 arm of the 1·2 chromosome. Concerning the ·1 arm of the **1·18** and the ·17 arm of the **2·17** chromosomes, assumptions have been made which are based on the appearance of the $2n$ + **1·18** and $2n$ + **2·17** tertiary plants. The ·1 arm of the **1·18** chromosome consists of approximately the distal two-thirds of the ·1 arm of the 1·2 chromosome. The ·17 arm of the **2·17** chromosome consists, distally, of the ·17 arm of the 17·18 chromosome plus, proximally, of the inner third of the ·1 arm adjacent to the insertion region. This is diagrammed in figure 1 where the 1·2 chromosome is represented by the solid line, the 17·18 chromosome by the dashed lined.

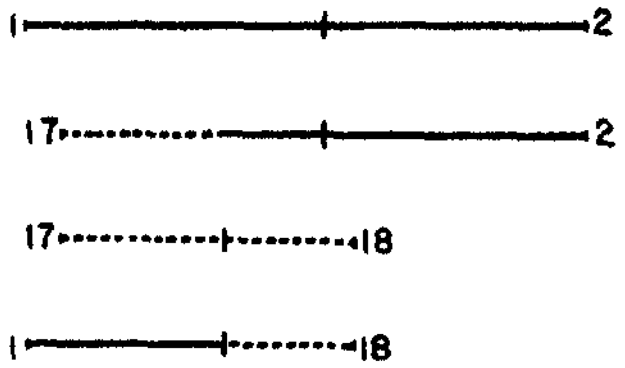


FIGURE 1

Lines representing approximate portions of the 1·2 and 17·18 chromosomes of *PT* 1 in the **1·18** and **2·17** chromosomes of *PT* 2. Solid line denotes 1·2 chromosome, dashed line the 17·18 chromosome.

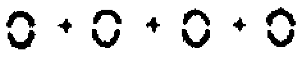
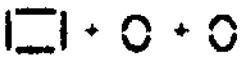
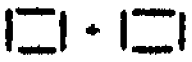
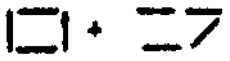
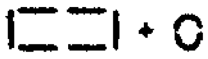
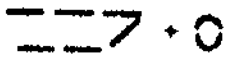
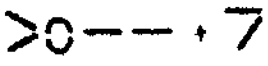
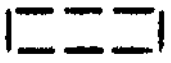
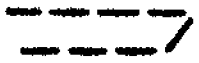
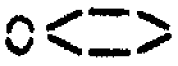
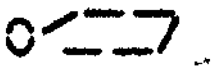
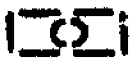
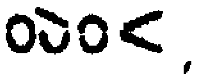
Tetraploid Prime Type 2.—In 1930[7] it was reported that a ⊙ 8 was seen in the $4n$ (*PT* 1 × *PT* 2) hybrid. At the Seventh International Genetics Congress[9] the ⊙ 8 and some other types of configuration were shown. It was suggested by Dr. F. W. Sansome that the relative frequency of the different possible types of configuration should be studied. Before this could be done, however, a new supply of $4n$ homozygous *PT* 2 plants was needed. By soaking seeds from $2n$ homozygous *PT* 2 plants in a solution of colchicine, such $4n$ plants were obtained. From the cross of $4n$ *PT* 1 × $4n$ *PT* 2 there were obtained hybrids which were grown to maturity. The present report is based on observations made on 45 *PMC* from ten of these plants.

Among the 48 chromosomes of $4n$ *PT* 1 there are four 1·2 and four 17·18 chromosomes; likewise for $4n$ *PT* 2 there are four **1·18** and four **2·17** chromosomes. In the $4n$ (*PT* 1 × *PT* 2) hybrid there are two each of the 1·2, **2·17**, **1·18** and 17·18 chromosomes. Whenever it is possible to identify these chromosomes, their arrangement at *M* I can be determined. Drawings of three of the many possible types of configuration involving these eight chromosomes are shown in figure 2, *B–D*. In *B* the sizes of the chromosomes in the ⊙ 8 are indicated. In *C* the four *L*-sized chromosomes form a ⊙ 4 to which is attached a *ch* 4 consisting of *m*-sized chromosomes. This configuration can be diagrammed as follows: $(1\cdot 2-\mathbf{2\cdot 17})_2^{10}$-17·18-**18·1**-**1·18**-18·17. In D there are two ⊙'s 4, each of which is (1·2-**2·17**-17·18-**18·1**).

In table 1 are listed the chromosomal arrangements seen in the 45 *PMC* referred to above. This number is admittedly too small for a study of rela-

TABLE 1

Types of Configuration Involving the 1·2, 1·18, 2·17 and 17·18 Chromosomes at Meiotic First Metaphase in 45 Pollen-Mother-Cells from Ten Tetraploid F₁ Hybrids (Prime Type 1 × Prime Type 2) of *Datura stramonium*

DIAGRAM OF CONFIGURATION	a	CHROMOSOMES IDENTIFIED AS	NO.	
1 	1	$(1\cdot2)_2 + (2\cdot17)_2 + (17\cdot18)_2 + (1\cdot18)_2$	1	$8_{IV} + 4_{II}$
2 	4	$(1\cdot2 - 2\cdot17)_2 + (1\cdot18)_2 + (17\cdot18)_2$	1	10_{IV}
		$(2\cdot17 - 17\cdot18)_2 + (1\cdot2)_2 + (1\cdot18)_2$	1	10_{IV}
		$(1\cdot18 - 18\cdot17)_2 + (1\cdot2)_2 + (2\cdot17)_2$	1	$9_{IV} + 2_{II}$
3 	3	$(1\cdot2-2\cdot17-17\cdot18-18\cdot1)+(1\cdot2-2\cdot17-17\cdot18-18\cdot1)$	1	10_{IV}
		$(1\cdot2-2\cdot17-17\cdot18-18\cdot1)+(1\cdot2-2\cdot17-17\cdot18-18\cdot1)$	2	$9_{IV} + 2_{II}$
		$(1\cdot2-2\cdot17-17\cdot18-18\cdot1)+(1\cdot2-2\cdot17-17\cdot18-18\cdot1)$	1	$7_{IV} + 6_{II}$
		$(1\cdot2 - 2\cdot17)_2 + (1\cdot18 - 18\cdot17)_2$	3	10_{IV}
4 	5	$(1\cdot2 - 2\cdot17)_2 + 1\cdot18 - 18\cdot17 - 17\cdot18 - 18\cdot1$	1	$9_{IV} + 2_{II}$
5 	4	$(2\cdot17 - 17\cdot18 - 18\cdot1)_2 + (1\cdot2)_2$	1	$7_{IV} + 6_{II}$
		$(18\cdot1 - 1\cdot2 - 2\cdot17)_2 + (17\cdot18)_2$	1	$8_{IV} + 4_{II}$
6 	24	Chain of 6 + $(17\cdot18)_2$	1	10_{IV}
7 	8	$1\cdot2 -$ $1\cdot2 - (2\cdot17)_2 - 17\cdot18 - 18\cdot1 + 1\cdot18 - 18\cdot17$	1	$9_{IV} + 2_{II}$
8 	5	Sequence not recorded	3	10_{IV}
		Sequence not recorded	4	$9_{IV} + 2_{II}$
		Sequence not recorded	3	$8_{IV} + 4_{II}$
		Sequence not recorded	1	$7_{IV} + 6_{II}$
9 	40	Sequence not recorded	2	10_{IV}
		$2\cdot1-1\cdot18-18\cdot17-17\cdot18-18\cdot1-1\cdot2-2\cdot17-17\cdot2$	1	10_{IV}
10 	8	$(1\cdot2)_2 - (2\cdot17 - 17\cdot18 - 18\cdot1)_2$	1	$9_{IV} + 2_{II}$
		$(1\cdot2)_2 - (2\cdot17 - 17\cdot18 - 18\cdot1)_2$	1	$8_{IV} + 4_{II}$
		$(2\cdot1)_2 - (1\cdot18 - 18\cdot17 - 17\cdot2)_2$	1	$9_{IV} + 2_{II}$
11 	8	$(1\cdot2)_2 - 2\cdot17 - 17\cdot18 - 18\cdot1 - 1\cdot18 - 18\cdot17 - 17\cdot2$	1	10_{IV}
		$(2\cdot1)_2 - 1\cdot18 - 18\cdot17 - 17\cdot2 - 2\cdot17 - 17\cdot18 - 18\cdot1$	1	$9_{IV} + 2_{II}$
		$(2\cdot17)_2 - 17\cdot18 - 18\cdot1 - 1\cdot2 - 2\cdot1 - 1\cdot18 - 18\cdot17$	1	$8_{IV} + 4_{II}$
		$(17\cdot18)_2 - 18\cdot1 - 1\cdot2 - 2\cdot17 - 17\cdot2 - 2\cdot1 - 1\cdot18$	1	$9_{IV} + 2_{II}$
12 	4	$(1\cdot2)_2 - (2\cdot17 - 17\cdot18 - 18\cdot1)_2$	1	$8_{IV} + 4_{II}$
13 	8	$2\cdot17 - \quad 18\cdot1$ $(1\cdot2)_2 - 2\cdot17 - (17\cdot18)_2 - 18\cdot1$	1	$9_{IV} + 2_{II}$
		Unresolved configurations of 8	2	10_{IV}
		Unresolved configurations of 8	8	$9_{IV} + 2_{II}$
		Unresolved configurations of 8	1	$7_{IV} + 6_{II}$

a THEORETICAL NUMBER OF ARRANGEMENTS.
b ASSOCIATIONS OF THE OTHER CHROMOSOMES.

tive frequencies of the many possible types of configuration. In fact, not all of the possible types were found. Nor do these 45 *PMC* represent a random sample because no *PMC* was included if one or more univalents were present. This extreme precaution was taken to make certain that the plants were true tetraploids, having four each of the following ten kinds of chromosomes: 3·4, 5·6, 7·8, 9·10, 11·12, 13·14, 15·16, 19·20, 21·22, 23·24 and two each of the 1·2, **1·18**, **2·17** and 17·18 chromosomes.

In column 1 of table 1 are listed in diagrammatic form thirteen types of configuration found in 39 *PMC*. In column 2 the theoretically possible number of arrangements in each of these types is given. Whenever the 1·2, **1·18**, **2·17** and 17·18 chromosomes could be identified, as was possible in 26 *PMC*, the sequences of the chromosomes in the configurations could be determined. They are listed in column 3. Further separations were made according to the number of quadrivalents and bivalents formed by the 40 other chromosomes. They are indicated in column 5 where the subscript $_{IV}$ denotes quadrivalents and the subscript $_{II}$ denotes *BV*. These latter data are included to show to what extent quadrivalents were formed in the same *PMC* under consideration. This tendency to form quadrivalents is consistent with the known behavior of other 4*n* Daturas.

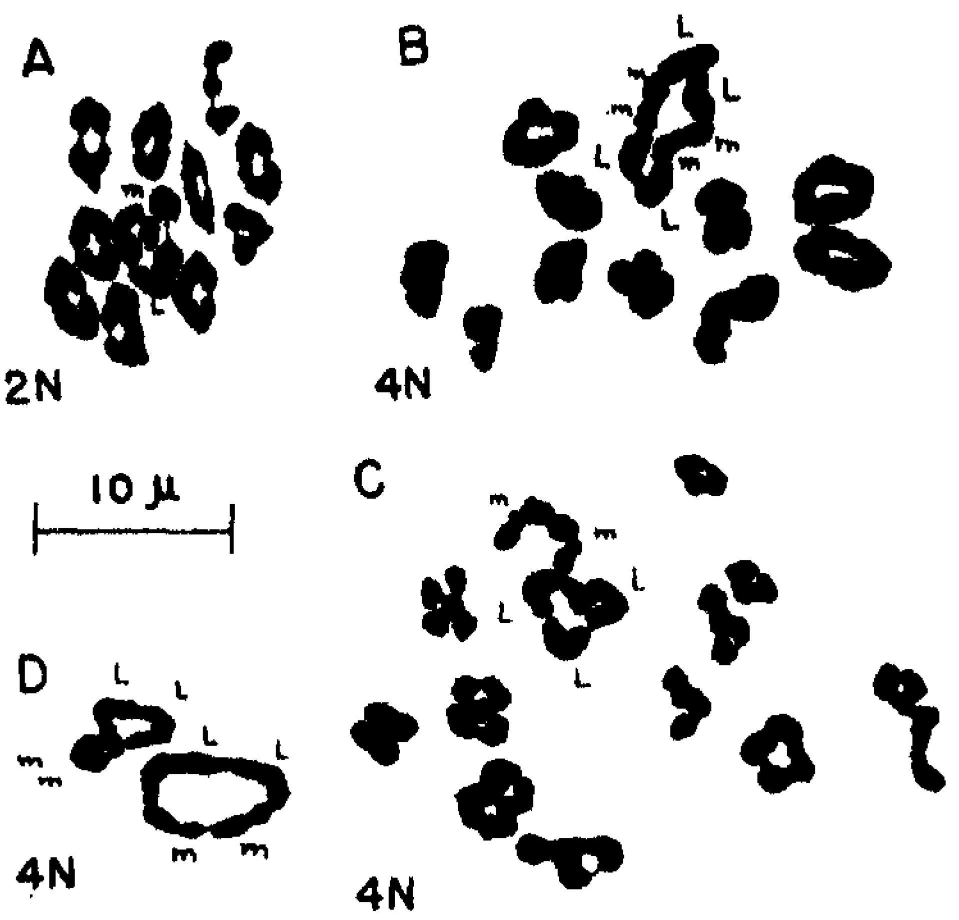


FIGURE 2

Chromosome associations at meiotic first metaphase in pollen-mother-cells of 2*n* and 4*n* (*PT* 1 × *PT* 2) plants of *Datura stramonium*. *A*. 2*n*: $\odot$ 4 (*LLmm*) + 10 *BV*. *B*. 4*n*: $\odot$ 8 (*LmmLLmmL*) + 10 quadrivalents. *C*. 4*n*: $\odot$ 4 (*LLLL*) with attached *ch* 4 —*mmmm* + 9 quadrivalents + 2 *BV*. *D*. 4*n*: two $\odot$'s 4 (*LLmm*); other chromosomes omitted.

Of the thirteen types listed, nine are distinct. Numbers 4, 6, 9 and 11 are modifications of Nos. 3, 5, 8 and 10, respectively. They form a series extending from four *B V* through a ⊙ 4 + 2 *B V*, two ⊙'s 4, ⊙ 6 + *B V* to a ⊙ 8. One quadruple attachment transforms the ⊙ 8 to a ⊙ 6 attached to a *B V*.

Discussion.—In the 2*n* (*PT* 1 × *PT* 2) hybrid, since there is only one each of 1·2, **1·18**, **2·17** and 17·18, a ⊙ 4 or *ch* 4 is the largest possible configuration and it is found regularly except when there are two pairs. In the 4*n* (*PT* 1 × *PT* 2) hybrid, however, there are two each of these chromosomes. Consequently one may find four *B V*. In addition many types of configurations which involve four, six or eight chromosomes are possible. Even in so modest a sample as 39 *PMC*, nine types plus four modifications were found.

Table 1 shows that practically two-thirds of the configurations involved all eight chromosomes. Moreover the ⊙ 8 was the type found most often (24 per cent). By including the *ch* 8 also the proportion was 31 per cent. Seven *PMC* (16 per cent) showed a ⊙ 6 or *ch* 6 attached to a *B V*. These were the commonest alternatives to a ⊙ 8 and involved one multiple attachment. Closely related is the ⊙ 6 or *ch* 6 plus a *B V*, where the *B V* is not attached to the other six chromosomes. If Nos. 5, 6, 10 and 11 are combined, their total equals 22 per cent.

Configurations involving four chromosomes were second in frequency. Two ⊙'s 4 each or a ⊙ 4 + *ch* 4 were found in 8 *PMC* or almost 18 per cent. One ⊙ 4 + 2 *B V* amounted to 7 per cent. Combining these two types gives a total of over 24 per cent.

Most of the configurations which involved four or more chromosomes were merely concatenations since they consisted solely of a series of attachments at the homologous ends of only *two* chromosomes. Theoretically three or four chromosomes could be attached together at their homologous ends. In 10 *PMC* there were 14 instances of such multiple attachments. Six triple attachments are listed under Nos. 7, 11 and 13 in table 1. Eight quadruple attachments are listed under Nos. 7, 10, 12 and 13. Three *PMC* showed more than one multiple attachment: No. 7 showed one triple and one quadruple, No. 12 showed two quadruple and No. 13 showed one triple and two quadruple attachments.

Seven *PMC* are classified under Nos. 10 and 11 in table 1. The multiple attachment involved a *B V* to which was attached a ⊙ 6 or *ch* 6. In five *PMC* the 1·2 chromosomes formed the *B V*, in one the **2·17**. Also in No. 12 the 1·2 chromosomes formed a *B V*. Possibly because of their large size the 1·2 and **2·17** chromosomes may form a *B V* more frequently than the *m*-sized ones.

It is to be regretted that the sequences of the chromosomes in the many ⊙'s 8 were not recorded with sufficient certainty to warrant publication.

In one of the ⊙ 8 + 10_{IV} *PMC* the sequence probably was (1·2-2·1-**1·18**-18·17-**17·2-2·17**-17·18-**18·1**). In one of the ⊙ 8 + 9_{IV} + 2_{II} *PMC* the sequence probably was (2·1-**1·18**-18·17-17·18-**18·1**-1·2-**2·17-17·2**). This same sequence is given for one of the *ch* 8 + 10_{IV} *PMC.*

In the 2*n* hybrid only one kind of ⊙ 4 can be formed, namely (1·2-**2·17**-17·18-**18·1**). In the 4*n* hybrid, however, in addition to this arrangement, other arrangements in a ⊙ 4 are possible because each of these chromosomes is present in duplicate. It is possible to have a ⊙ 4 which consists of two of each of two kinds of chromosomes. These possibilities are as follows: (1·2-**2·17**)$_2$, (**1·18**-18·17)$_2$, (**2·17**-17·18)$_2$ and (2·1-**1·18**)$_2$. This last kind was not found at all and (**2·17**-17·18)$_2$ was found only once in a ⊙ 4 + 2 *BV* type of configuration.

There are three possible arrangements for two configurations of four each as follows: (2·1-**1·18**-18·17-**17·2**) + (2·1-**1·18**-18·17-**17·2**), (1·2-**2·17**)$_2$ + (**1·18**-18·17)$_2$ and (2·1-**1·18**)$_2$ + (**2·17**-17·18)$_2$. In seven *PMC* two ⊙'s 4 were found and in another a ⊙ 4 + *ch* 4. If these two types are considered together, it is seen that the first two arrangements were identified and found with equal frequency. However, the third possible arrangement was not identified.

Summary.—In tetraploid *Datura* hybrids between prime type 1 and prime type 2 there are two each of the 1·2 and 17·18 chromosomes from *PT* 1 and of the **1·18** and **2·17** chromosomes from *PT* 2 and four each of the other chromosomes.

In 45 *PMC* from ten plants the configurations formed by associations among the 1·2, **1·18**, **2·17** and 17·18 chromosomes at *M* I were studied. Two-thirds of the configurations involved all eight chromosomes.

In 39 *PMC* the configurations could be diagrammed. Nine distinct types plus four modifications were found. The series extended from four *BV* through ⊙ 4 + 2 *BV*, two ⊙'s 4, ⊙ 6 + *BV* to a ⊙ 8. In 24 per cent of the *PMC* the ⊙ 8 was found. A common modification was a *ch* 8. A ⊙ 6 or *ch* 6 attached to a *BV* also was seen frequently. Two configurations of four each were the next commonest types.

Usually a configuration which involved four or more chromosomes consisted solely of a series of attachments at the homologous ends of only two chromosomes (concatenation). Only fourteen instances of attachment at the homologous ends of three or four chromosomes were found.

The 1·2 chromosomes formed the *BV* in five of the seven *PMC* in which a ⊙ 6 or *ch* 6 was attached to a *BV*.

In 2*n* hybrids between *PT* 1 and *PT* 2 only one kind of chromosome arrangement is possible in the ⊙ 4 (1·2-**2·17**-17·18-**18·1**). In 4*n* hybrids three additional arrangements in a ⊙ 4 were found, namely (1·2-**2·17**)$_2$, (**1·18**-18·17)$_2$ and (**2·17**-17·18)$_2$.

[*] This work was done at the Carnegie Institution of Washington, Department of Genetics, Cold Spring Harbor, New York, while the author was associated with Dr. A. F. Blakeslee.

[1] Bergner, A. D., Satina, S., and Blakeslee, A. F., these PROCEEDINGS, 19, 103–115 (1933).

[2] Blakeslee, A. F., *Univ. Pennsylvania Bicentennial Conf.* In Cytology, genetics, and evolution, 37–46 (1941).

[3] Blakeslee, A. F., Bergner, A. D., and Avery, A. G., *Cytologia*, Fujii Jub.Vol., 1070–1093 (1937).

[4] Numbers representing chromosomes other than those of *PT* 1 are shown in boldface type.

[5] ⊙ denotes circle.

[6] Parentheses enclose chromosomes within a ⊙.

[7] Blakeslee, A. F., *Carnegie Inst. Wash. Year Book*, 29, 39 (1930).

[8] Satina, S., Bergner, A. D., and Blakeslee, A. F., *Amer. Jour. Bot.*, 28, 383–390 (1941).

[9] Bergner, A. D., *Proc. 7th. Intern. Genet. Congress*, 63–64 (1939).

[10] Parentheses with subscript $_2$ denote that ⊙ consists of two of each kind of chromosome included within the parentheses.

ON MECHANICAL SELF-EXCITED OSCILLATIONS

BY N. MINORSKY

DAVID TAYLOR MODEL BASIN, NAVY DEPARTMENT

Communicated August 28, 1944

1. *Introduction.*—Self-excited oscillations, particularly electrical ones, form the object of numerous studies in recent years. Mechanical oscillations of this kind have been less explored although they are frequently observed in practice.

The object of this note is to describe a typical phenomenon of this kind observed during experimental work on the antirolling stabilization of ships by the activated tanks method. The problem of the antirolling stabilization of ships itself does not form the object of the note and only a few words will be sufficient to explain the principal details of these experiments.

2. *Experimental Arrangement and Principal Data.*—The principal features of ship stabilization by tanks can be conveniently studied both theoretically and experimentally by means of two coupled pendula of which one P is a rigid physical pendulum and the other P' is a pendulum consisting of a liquid (water, oil, etc.) filling a U-tube connected to two tanks and secured to the rigid pendulum P so as to obtain an oscillation coplanar with the latter.

In the experiments described here the activation of the ballast was accomplished by means of an axial variable pitch pump inserted in the U-channel of the liquid pendulum. The impeller pump was kept running at a constant speed and the activation was accomplished by controlling the blade angle α of the runner by a control equipment responsive to the angular motion of P.

Let θ be the angle of P and φ the relative angle formed by the levels of P' with respect to P. The rate of flow of the ballast is then proportional to $\dot{\varphi}$. The following experimental results will be of importance.

1. It was found that the roll-quenching efficiency is greatest when the following *condition of control* is fulfilled

$$\alpha = \lambda\ddot{\theta} \tag{1}$$

that is, when the blade angle α is made to vary in proportion to the angular acceleration $\ddot{\theta}$ of the rigid pendulum. The coefficient of proportionality λ could be adjusted by changing the amplification factor of an electronic circuit controlling the blade angle α.

2. For a steady harmonic actuation of the ballast by the impeller pump it is observed that there exists the following relation:

$$\dot{\varphi} = \epsilon\ddot{\theta} \tag{2}$$

between the rate of flow $\dot{\varphi}$ and the angular acceleration $\ddot{\theta}$; this means in view of (1) that the rate of flow $\dot{\varphi}$ and the blade angle α are approximately proportional to each other or,

$$\alpha = \lambda/\dot{\epsilon} = \gamma\dot{\varphi} \tag{3}$$

the coefficient ϵ changes somewhat according to conditions.

3. Moreover, from numerous tests, it has been ascertained that the force imparted by the impeller pump to the liquid column increases more or less linearly with α for small values of α and increases less than in proportion for larger values of α. The moment M of the force created by the impeller and applied to P' can be, therefore, approximated by a relation

$$M = M_1\alpha - M_3\alpha^3 \tag{4}$$

where M_1 and M_3 are independent of α. The form of the curves approximated by this expression varies somewhat for different conditions.

3. Phenomenon of Self Excitation.—It has been ascertained that when the coefficient λ of amplification is made large enough, a self-excited oscillation of relatively high frequency sets in and superimposes itself on the normal stabilizing action. If the coefficient λ continues to increase further, self-excited oscillation exhibits the presence of harmonics. If, however, λ is decreased gradually, at a certain point $\lambda = \lambda_0$ the self-excited oscillation disappears. During the experiments the parasitic fluttering of the blades had a frequency nearly five times that of the oscillatory process which the equipment is intended to quench.

The interesting feature of the phenomenon lies in the fact that it may originate also when the pendulum P is almost at rest. In fact it is sufficient to disturb the pendulum at rest very little, e.g., by touching it slightly, to be able to release violent surges of water in the tanks with the corresponding blade angle oscillations, which persist indefinitely. The pendulum P also begins to oscillate with the same high frequency but with an amplitude of a fraction of one degree whereas the oscillation of P' is generally very appreciable and with strong amplification may reach the value of about 20 to 25 degrees. In what follows we shall investigate the problem in this particular case which is relatively simple.

4. *Differential Equations.*—The fact that during the self excitation from rest the amplitude of motion in the θ coördinate is very small, makes it possible to neglect the effect of the passive component of coupling between P and P' since this component depends on the existence of a finite oscillation in the θ degree of freedom, and consider only the active component of coupling, i.e., the action of the moment M due to the pump on P' and the reaction of this moment on P.

The simplified differential equations of the system in this case are:

$$j\ddot{\varphi} + k\dot{\varphi} + c\varphi = M; \qquad I\ddot{\theta} + K\dot{\theta} + C\theta = -M \tag{5}$$

where j, k and c for the liquid pendulum P' and I, K and C for the rigid pendulum P are certain constants, the physical significance of which is sufficiently obvious. The left-hand terms of these equations are linearized somewhat which has no further influence on what follows. The neglect of the passive component of the coupling as just mentioned is in the fact that the first equation contains only the relative coördinate φ instead of the absolute one $\varphi - \theta$ since in this particular case θ is very small.

Equations (5) are thus written down with the knowledge of the principal features of the phenomenon and our purpose will be to fit these features into the form of solutions we are going to obtain. Using the equations (4) and (3) we obtain

$$j\ddot{\varphi} + (k - m_1)\dot{\varphi} + m_3\dot{\varphi}^3 + c\varphi = 0 \tag{6}$$

$$I\ddot{\theta} + K\dot{\theta} + C\theta = -m_1\dot{\varphi} + m_3\dot{\varphi}^3 \tag{7}$$

where $$m_1 = M_1\lambda/\epsilon \qquad \text{and} \qquad m_3 = M_3\lambda^3/\epsilon^3. \tag{8}$$

It is seen that owing to the neglect of the passive component of coupling equations (6) and (7) do not form actually a coupled system since equation (6) contains only φ. Hence, if we succeed in solving the non-linear equation, φ will appear as a known function in (7) which will permit determining θ and through the equation (1) of control we shall be able to establish the condition of coupling.

Thus, the principal interest centers on the equation (6). It is noted that for a sufficiently small value of λ, the coefficient m_1 is also small so that $k - m_1 > 0$. In such a case it is apparent that no self excitation is possible in the φ degree of freedom since equation (6) shows that the liquid pendulum P' has a purely dissipative damping of a combined velocity, and velocity-cube type. The really interesting case arises when the parameter λ is increased so as to make the coefficient $k - m_1 < 0$ which case we shall consider from now on. Physically this means that the energy imparted by the blades outweighs the dissipation of energy caused by the velocity damping $k\dot{\varphi}$.

In order to bring the problem within the scope of non-linear mechanics we shall assume that the damping terms are relatively small in comparison with the inertial $(j\ddot{\varphi})$ and stability $(c\varphi)$ terms which is sufficiently justified by the actual order of magnitude of these quantities. This permits writing

$$m_1 - k = \mu s; \qquad m_3 = \mu q$$

where μ is a small parameter $(\mu > 0)$ and $s > 0$ and $q > 0$ are certain constants of the order of j and c. With these conventions equation (6) can be written

$$j\ddot{\varphi} + c\varphi + \mu(-s\dot{\varphi} + q\dot{\varphi}^3) = 0. \tag{9}$$

This equation is of the Rayleigh type. By differentiating it becomes of the Van der Pol type in $\dot{\varphi}$ and thus may be shown[1] to have the approximate solution

$$\varphi = \frac{1}{\omega_0}\sqrt{\frac{4(m_1 - k)}{3m_3}}\cos(\omega_0 t + \psi_0)$$

$$\dot{\varphi} = -\sqrt{\frac{4(m_1 - k)}{3m_3}}\sin(\omega_0 t + \psi_0). \tag{10}$$

Substituting this value for $\dot{\varphi}$ into equation (7) one easily obtains θ since this equation is linear. Since the term $m_3\dot{\varphi}^3$ contains the third harmonic it is apparent that the oscillation of P will also contain that harmonic and in view of equations (1) and (3) this harmonic will also appear in the φ oscillation. The amplitudes of these harmonics are small if the ratio m_3/m_1 is small as we have assumed. We shall limit this study to the first approximation only assuming that $m_3 = 0$ in equation (7). Since the phase ψ_0 is clearly of no interest here we can write

$$I\ddot{\theta} + K\dot{\theta} + C\theta = -m_1\dot{\varphi} = m_1\sqrt{\frac{4(m_1 - k)}{3m_3}}e^{i\omega_0 t} = Ae^{i\omega_0 t}$$

where $A = m_1\sqrt{\frac{4(m_1 - k)}{3m_3}}$; whence the approximate expression for the amplitude θ_0 is

$$\theta_0 = \frac{m_1}{I\omega_0^2}\sqrt{\frac{4(m_1 - k)}{3m_3}}$$

and the amplitude of the angular acceleration is then

$$\ddot{\theta}_0 = \frac{m_1}{I}\sqrt{\frac{4(m_1 - k)}{3m_3}}. \tag{11}$$

The blade angle is given by equation (1)

$$\alpha_0 = \lambda\ddot{\theta}_0.$$

It is seen that the blade angle generally increases with the amplification factor but not exactly in proportion to λ since by equations (8) both m_1 and m_3 are functions of λ. It is generally observed that the blade angle α increases with the amplification factor but the observation is handicapped owing to the presence of harmonics which are left out in this study.

6. *Critical Value of Parameter.*—We shall now attempt to establish analytically the principal feature of the phenomenon mentioned at the beginning of Section 3, namely, a rather sudden appearance of self excitation at a certain value $\lambda = \lambda_0$ of the parameter when it is increased from small values as well as its sudden disappearance at the same value $\lambda = \lambda_0$, when λ is gradually decreased from larger values.

Although this fact has been discussed to some extent in Section 4 we shall analyze it now from a more general standpoint resulting from the theory of Poincaré-Liapounoff.[2] Only equation (6) will interest us here because equation (7) does not present any difficulty.

Putting $m_1 = \lambda a_1$ and $m_3 = \lambda^3 a_3$ where a_1 and a_3 are certain positive constants at least within the range of the experiment as follows from (8) we write equation (6) in the form

$$j\ddot{\varphi} + (k - \lambda a_1)\dot{\varphi} + m_3\dot{\varphi}^3 + c\varphi = 0. \tag{6a}$$

This equation is equivalent to the following system of differential equations

$$\begin{aligned} \dot{\varphi} &= y \\ \dot{y} &= -g(\lambda)y - \omega_0^2\varphi - m_3'y^3 \end{aligned} \tag{12}$$

where

$$g(\lambda) = (k - \lambda a_1)/j \text{ and } m_3' = m_3/j.$$

We shall investigate the conditions of stability of the liquid pendulum when at rest. Liapounoff has shown that for this purpose it is legitimate to neglect the non-linear term in the second equation (12) which gives

$$\dot{\varphi} = y$$
$$\dot{y} = -g(\lambda)y - \omega_0^2\varphi. \qquad (13)$$

We can now apply the classical method of Poincaré which consists in investigating the singularities of the system (13) as a function of the parameter. In this case there is clearly one singular point of interest, viz., $\varphi = y = 0$. The characteristic equation is

$$S^2 + g(\lambda)S + \omega_0^2 = 0 \qquad (14)$$

and its roots are

$$S_{1,2} = -g(\lambda)/2 \pm \sqrt{\frac{g^2(\lambda)}{4} - \omega_0^2}. \qquad (15)$$

There are four possibilities according to the sign of $g(\lambda)/2$ and of the quantity under the square root.

If $\left|\frac{g(\lambda)}{2}\right| > \omega_0$ the roots are real and of the same sign which corresponds to a nodal point of the system. This case is to be discarded because the experiment always shows that a self-excited oscillation builds itself up in a series of swings and never reaches a steady state aperiodically which would require a pump of an enormous power. Hence, we have to consider the case when $\left|\frac{g(\lambda)}{2}\right| < \omega_0$ in which case the roots are conjugate complex and the singularity $\varphi = y = 0$ is a focal point, stable if $g(\lambda) > 0$ and unstable if $g(\lambda) < 0$. The value $\lambda = \lambda_0$ for which $g(\lambda_0) = 0$ and which separates the stable focal points from the unstable ones is thus the critical or *bifurcation* value of the parameter λ (Poincaré).

Since $g(\lambda)$ is monotone decreasing with λ increasing and is positive for small λ, it follows that by gradually increasing the value of the parameter λ the original stable focal point for $\lambda \leqslant \lambda_0$ becomes unstable for $\lambda \geqslant \lambda_0$.

From the theory of Poincaré it follows that for the critical value $\lambda = \lambda_0$ for which a stable singularity becomes unstable there may appear at least one stable *limit cycle* which means a stationary oscillation. The proof of the existence of a limit cycle is rather laborious and requires the investigation of a series expansion satisfying the non-linear system (12). This need not be done here since we know that the self-excited oscillation exists for $\lambda \geqslant \lambda_0$ and the theory of the first approximation gives the answer in a simple manner.

7. *Conclusion.*—It is seen that the principal features of the phenomenon in question are consistently explained on the basis of non-linear mechanics. In fact, if one had to follow purely linear methods it would be impossible even to predict the existence of a phenomenon of this nature.

[1] Kryloff, N., and Bogoliuboff, N., *Introduction á la Mécanique non linéaire*, Kieff, 1937; also free English translation by S. Lefschetz, Princeton University Press, 1943.

[2] Poincaré, H., "Sur l'équilibre d'une masse fluide," *Acta Mathematica*, Vol. 7; *Figures d'équilibre d'une masse fluide*, Paris, 1903. Liapounoff, A., Problème général de la stabilité du mouvement. *Ann. de la Fac, d. Sc. de Toulouse*, **9**, 203 (1907). Andronow, A., and Chaikin, S., *Theory of Oscillations*, Moscow, 1937.

FIRST PROOF THAT THE MERSENNE NUMBER M_{157} IS COMPOSITE

By H. S. Uhler

Sloane Physics Laboratory, Yale University

Communicated September 9, 1944

The form of a Mersenne number is $M_p = 2^p - 1$ where p is prime. Just three hundred years ago Mersenne published, in effect, the statement that the only values of p not greater than 257 which make M_p prime are 2, 3, 5, 7, 13, 17, 19, 31, 67, 127 and 257. The prime or composite character of all of the 55 numbers included under this conjecture, except the six corresponding, respectively, to $p = 157$, 167, 193, 199, 227 and 229, had been investigated prior to the year 1935. Contrary to Mersenne's surmise it was found that for $p = 67$ and 257 M_p is composite, and that M_p is prime for $p = 61$, 89 and 107. The data presented above have been derived from a comprehensive paper by R. C. Archibald.[1]

The present contribution marks the first fruits of a friendly suggestion made last year by Professor Archibald and seconded by Professor D. H. Lehmer that the author turn his attention to problems of factorization in general, beginning with the special problem of the character of the six heretofore uninvestigated Mersenne numbers.

The modern technique of this problem is based explicitly upon the following theorem, discovered by E. Lucas and clarified by D. H. Lehmer,[2] namely: "*The number $N = 2^n - 1$, where n is an odd prime, is a prime if, and only if, N divides the $(n - 1)$-st term of the series,*

$$S_1 = 4, \qquad S_2 = 14, \qquad S_3 = 194, \ldots, S_k, \ldots,$$

where $S_k = S_{k-1}^2 - 2$."

Since the number M_{157}, which is the subject of the present study, equals 182 68770 46663 62864 77546 06040 89535 37745 69915 67871 and since the above theorem requires that this 48-digit number be used as a divisor 149 times (8 to 156, inclusive) it should be obvious that the prospective investigator would focus attention upon this aspect of the computation. The procedure followed by R. E. Powers[3] in showing that M_{241} is composite appealed at first to the author and hence it was used as far as the 34th member of the Lucasian sequence. At this stage Powers' technique was abandoned because it had proved to involve an inordinate amount of writing and mental arithmetic and especially because it did not make use of a computing machine exclusively.

The patent idea of substituting for division multiplication by the reciprocal of the divisor overcame all of the objections entertained by the writer against all of the schemes for direct machine-division which he had read in the literature. Consider $s_k' = N \cdot q_k + r_k$. Then $(1/N) \cdot s_k' = q_k + r_k/N$. Hence if we multiply any term $(> N)$ of a congruently equivalent Lucasian sequence by $1/N$ the succession of figures to the left of the decimal point will represent the integer q_k, while the digits to the right of the decimal point will constitute an approximation to the incommensurable decimal fraction r_k/N. Theoretically r_k may be obtained by multiplying r_k/N by N without evaluating a single q_k at any stage of the work. This procedure was not followed because it involves a greater number of figures than the alternative process of multiplying q_k by N and then subtracting $N \cdot q_k$ from s_k', that is, defining r_k as $s_k' - N \cdot q_k$.

The following details of the actual computation should make the matter perfectly clear. Ribbon A of cross-ruled paper has transcribed in nonads along the lower edge a sufficient number of figures of the reciprocal of $N(M_{157})$ thus: $1/N = 0 \cdot$(47 zeros) 547382212 626881668 329581868 472623488 781352321 048510246 · · ·. Similarly the bottom line of ribbon B reads: $N =$ 182687704 666362864 775460604 089535377 456991567 871.000000. Ribbons A and B do not have to be changed throughout the work from $k = 8$ to $k = 156$. On the other hand, each value of k requires three ribbons of its own. Along the top line of a longer ribbon of cross-ruled paper is written, for example, the value of r_{155} as 744260418 487541370 275270351 147668949 680082908 44.0000000. This number was machine-squared by D. N. Lehmer's[4] algorithm and recorded along the lowest line of the same ribbon. After subtracting 2 this line read $s_{156}' =$ 553923570 527250212 804427560 946230012 742643893 708358589 871125964 130378182 986908821 472209423 2334. The undiminished squared number was carefully checked with the moduli $10^3 + 1$, $10^4 + 1$ and $10^7 + 1$. Next ribbon A was laid on the longer ribbon so that the nonads of $1/N$ were juxtaposed above and in exact register with the first six nonads of s_{156}'. The multiplication of s_{156}' by $1/N$ was then effected and the resulting value of the integer q_{156} presented itself on the bottom line of a fourth strip of paper. This short act of multiplying was always repeated as a preliminary check on its accuracy. Actually $q_{156} =$ 303207909 661388759 507868593 612664030 922503212 07. Finally ribbon B was aligned with the nonads of q_{156} and a portion of the product $N \cdot q_{156}$ was recorded on the second line from the bottom of a fifth strip of cross-ruled paper. This multiplication was performed only on the nonads from the fifth to the eleventh, inclusive, since the first four sums-of-products would have reproduced a part of s_{156}' and not of r_{156}. (All of the nonads after the sixth were composed entirely of supplied zeros both for N and q_{156}.) The last figures of s_{156}' were copied on the third line from the bottom of the fifth strip so that by forming the difference between the second and third lines the required value of r_{156} came out on the bottom line. This remainder was found on Aug. 11, 1944, to be 118 57508 80382 71696 98184 73569 85091 23773 18030 92037. Since this residue is not zero it follows that M_{157} is composite and incidentally that M_{127} still retains the position of being the largest known prime number. For every value of k from 8 to 156 the numbers on the three corresponding strips were found to satisfy the relation $s_k' = N \cdot q_k + r_k$ for each of the moduli $10^5 + 1$, $10^4 + 1$ and $10^7 + 1$. The author desires to announce that he has already begun to investigate M_{167}.

[1] Archibald, R. C., *Scripta Mathematica*, **3**, 112–119 (1935).

[2] Lehmer, D. H., *Jour. London Math. Soc.*, **9–10**, 162–165 (1934–1935).

[3] Powers, R. E., *Bull. Amer. Math. Soc.*, **40**, 883 (1934).

[4] Lehmer, D. N., *Amer. Math. Monthly*, **30**, 67, 68 (1923).

ON THE STABILITY OF TWO-DIMENSIONAL PARALLEL FLOWS*

BY C. C. LIN

GUGGENHEIM LABORATORY, CALIFORNIA INSTITUTE OF TECHNOLOGY

Communicated August 24, 1944

1. Introduction.—Heisenberg's[1] remarkable contribution to the hydrodynamic stability of two-dimensional parallel flows has not been favorably accepted and properly appreciated, because his paper is not completely free from obscure points. Nor has the work of Tietjens,[2] Tollmien[3] and Schlichting[4] been properly estimated. As a result, the theory to account for the instability of laminar flow at high Reynolds numbers has become very confused, and its further development has been very much retarded. Various authors suggest that it is necessary (1) to consider disturbances of finite amplitudes, (2) to include the effect of compressibility or even (3) to modify the Navier-Stokes equations. The present situation of our knowledge may be seen from the general lectures given by G. I. Taylor[5] and J. L. Synge.[6]

Recently the present writer carried out some investigations in an attempt to clarify the situation. The theory of Heisenberg was critically examined, somewhat modified and further developed. These developments were made with particular emphasis on the general criteria of instability and their underlying physical mechanism, both in an inviscid fluid and in a real fluid. According to the general criterion obtained, the plane Poiseuille flow and the Blasius flow are both unstable at sufficiently high Reynolds numbers.[7] Detailed numerical calculations were also worked out in these cases. The present article gives a brief account of the essential results obtained. A detailed treatment will be published elsewhere.

2. *The Equation of Orr and Sommerfeld.*—In studying the stability of two-dimensional parallel flows, we may restrict ourselves to two-dimensional disturbances, according to an investigation of Squire.[8] A two-dimensional periodic disturbance in a field of flow with main velocity $w(y)$ parallel to the x-axis may be represented by the stream function $\psi = \varphi(y)e^{i\alpha(x-ct)}$, and the linearized differential equation for $\varphi(y)$ is

$$(w - c)(\varphi'' - \alpha^2\varphi) - w''\varphi = -\frac{i}{\alpha R}(\varphi^{iv} - 2\alpha^2\varphi'' + \alpha^4\varphi). \quad (2.1)$$

All the quantities involved are dimensionless: R is the Reynolds number of the main flow, and α, c are constants describing the nature of the disturbance. We shall regard α as real so that $2\pi/\alpha$ is the wave-length of the disturbance. The real part of c gives the phase velocity while its imaginary part (multiplied by α) gives the index of damping or magnification. The function $w(y)$ is regarded as defined for all complex values of y by analytic extension. Physical considerations then lead to certain homogeneous boundary conditions to be satisfied by $\varphi(y)$ at two real points y_1 and y_2 of the complex y-plane, which represent the coördinates of certain layers in the field of flow. There are altogether four boundary conditions. Hence, we are led to an eigenvalue problem, and a relation of the type

$$F(\alpha, c, \alpha R) = 0 \quad \text{or} \quad c = c(\alpha, R) \quad (2.2)$$

holds, where F is indeed an entire function of the parameters. Since stability, neutral stability or instability of the motion correspond to the real part of $-i\alpha c$ less than, equal to or greater than zero, the problem is to determine the imaginary part $c_i(\alpha, R)$ of $c(\alpha, R)$ for a set of real values of α and R.

The actual investigation depends upon the solution of (2.1) by the method of successive approximations. Two methods are used. First, we introduce the variable

$$\eta = (y - y_0)/\epsilon, \quad (2.3)$$

where $w(y_0) = c$, $\epsilon = (\alpha R)^{-1/3}$, and $w'(y_0)$ is positive for real values of c. In this way, we obtain a fundamental system of four solutions $\varphi(y)$ in the form

$$\varphi(y) = \sum_{n=0}^{\infty} \epsilon^n \chi^{(n)}(\eta), \tag{2.4}$$

where $\chi^{(n)}(\eta)$ are explicitly expressible in terms of the Hankel functions $H_{1/3}^{(1),(2)}[{}^2/_3(i\alpha_0\eta)^{3/2}]$, $\alpha_0 = \{w'(y_0)\}^{1/3}$. Indeed, the series (2.4) can be shown to be convergent for the range of variables and parameters concerned. Secondly, we may put

$$\varphi(y) = \sum_{n=0}^{\infty} (\alpha R)^{-n}\varphi^{(n)}(y), \tag{2.5}$$

which gives us two asymptotic solutions. The initial approximation of (2.5) satisfies the inviscid equation

$$(w - c)(\varphi^{(0)''} - \alpha^2\varphi^{(0)}) - w''\varphi^{(0)} = 0, \tag{2.6}$$

which can be solved by expanding $\varphi^{(0)}$ in power series of α^2, convergent for all values of α^2. Two other asymptotic solutions are obtained in the form

$$\varphi(y) = \exp\{\textstyle\int g dy\}, \qquad g = (\alpha R)^{1/2}g_0 + g_1 + (\alpha R)^{-1/2}g_2 + \ldots. \tag{2.7}$$

Both methods of solution are essentially due to Heisenberg,[1] and Tollmien followed the same way. Heisenberg also carried out some approximate calculation for the case of channel flow, while Tollmien studied the case of the boundary layer. Criticism of Heisenberg's work is usually associated with the multiple-valued nature of the asymptotic solutions and the convergence of the power series in α^2 used for solving (2.6). Detailed investigations justify Heisenberg's treatment. In fact, the asymptotic solutions hold only for certain regions of the y-plane to be determined by comparison with the asymptotic expansions of (2.4). Indeed, from the properties of the Hankel functions, it can be easily shown that under the restriction

$$-\frac{7\pi}{6} < \arg(\alpha_0\eta) < \frac{\pi}{6}, \tag{2.8}$$

the asymptotic expansions do not suffer any jump (which may be called the Stokes phenomenon or the Uebergangs-substitution). Fortunately, such a restriction does not exclude the existence of a connected region enclosing the points y_1 and y_2 where our boundary conditions are to be satisfied. The eigenvalue problem can therefore be conveniently treated with the help of the asymptotic solutions. However, the region of validity of the asymptotic expressions does not always include all the points of the real segment (y_1, y_2). As a matter of fact, the asymptotic solutions hold throughout the real interval (y_1, y_2) only when $c_i > 0$. For $c_i < 0$, there are two points on the real axis where the Stokes jump of the asymptotic solutions occur. These two points coalesce into one critical point when $c_i = 0$.

Physically, the failure of the asymptotic solutions at some points on the real axis (and therefore the failure of $\varphi^{(0)}$ as a first approximation) in the case of neutral ($c_i = 0$) and damped oscillations ($c_i < 0$) indicates the presence of the effect of viscosity at the corresponding layers inside the fluid. This was noted by several authors in the cases of neutral stability. However, the existence of *two inner viscous layers* for the damped disturbance does not seem to have been noticed before. As we shall see in Section 3, it has an important consequence upon the well-known criteria of Rayleigh[9] and Tollmien.[10]

3. *General Criteria of Stability and the Curve of Neutral Stability.*—In the limit of infinite Reynolds numbers, it can be shown by careful mathematical investigation that the problem can be treated by finding a solution of (2.6) satisfying two boundary conditions. Rayleigh[9] and Tollmien[10] have shown that the necessary and sufficient condition for the existence of neutral or self-excited disturbances is that the velocity curve has a point of inflection in the interval (y_1, y_2). The necessary condition is general, while the sufficient condition has been proved by Tollmien only for velocity distributions of the symmetrical type or of the boundary-layer type. In fact, an example has been shown by the present writer where a point of inflexion exists in the velocity curve, but a neutral disturbance does not exist.

The authors mentioned above also concluded that damped disturbances are excluded if a point of inflexion does not exist. However, their proof assumes the analytical nature of the solution $\varphi^{(0)}(y)$ along the real axis, which is true for self-excited disturbances, but not for damped disturbances. Previous authors regarded a damped disturbance as the complex conjugate of a self-excited disturbance. This is in contradiction with the above-mentioned drastic difference of nature of those solutions along the real axis. Careful investigation reveals that the damped or self-excited nature of a solution is not changed by taking the complex conjugate. *Therefore damped disturbances are not excluded by the condition* $w''(y) \neq 0$ (at least not according to Rayleigh's proof).

Tollmien's proof to establish the existence of a point of inflexion in the velocity curve as sufficient to insure instability is rather cumbersome, and he has to assume that w''' does not vanish at the point of inflexion. These points have now been improved by a more careful consideration of the analytical nature of (2.6).

Having settled the question of hydrodynamic stability at infinite Reynolds numbers, the author studied the stability in a real fluid by trying to find out how the foregoing results should be modified for large but finite values of the Reynolds number. One such result has already been obtained by Heisenberg,[1] whose conclusion is as follows. If a velocity distribution allows an inviscid neutral disturbance with finite wave-length and nonvanishing phase velocity, the disturbance with the same wave-length is unstable in the real fluid when the Reynolds number is sufficiently large. In the present investigation, the proof has been formulated in a slightly more satisfactory manner.

Since a point of inflection in the velocity curve is sufficient to insure instability only for velocity distributions of the symmetrical and boundary-layer types, the author has also made more attempts in these cases. Indeed, it has been shown that *in these cases the flow is always unstable for sufficiently large Reynolds numbers*, whether the velocity curve has a point of inflection or not. The curve of neutral stability

$$c_i(\alpha, R) = 0 \qquad (3.1)$$

may belong to either of the types shown in figure 1. When the velocity curve has no point of inflection, the two asymptotic branches of the curve have the common asymptote $\alpha = 0$ (Fig. 1 *a*). When there is a point of inflection, one branch has the asymptote $\alpha = 0$, while the other has the asymptote $\alpha = \alpha_s > 0$ (Fig. 1 *b*). These results are in agreement with those of Rayleigh, Tollmien and Heisenberg cited above.

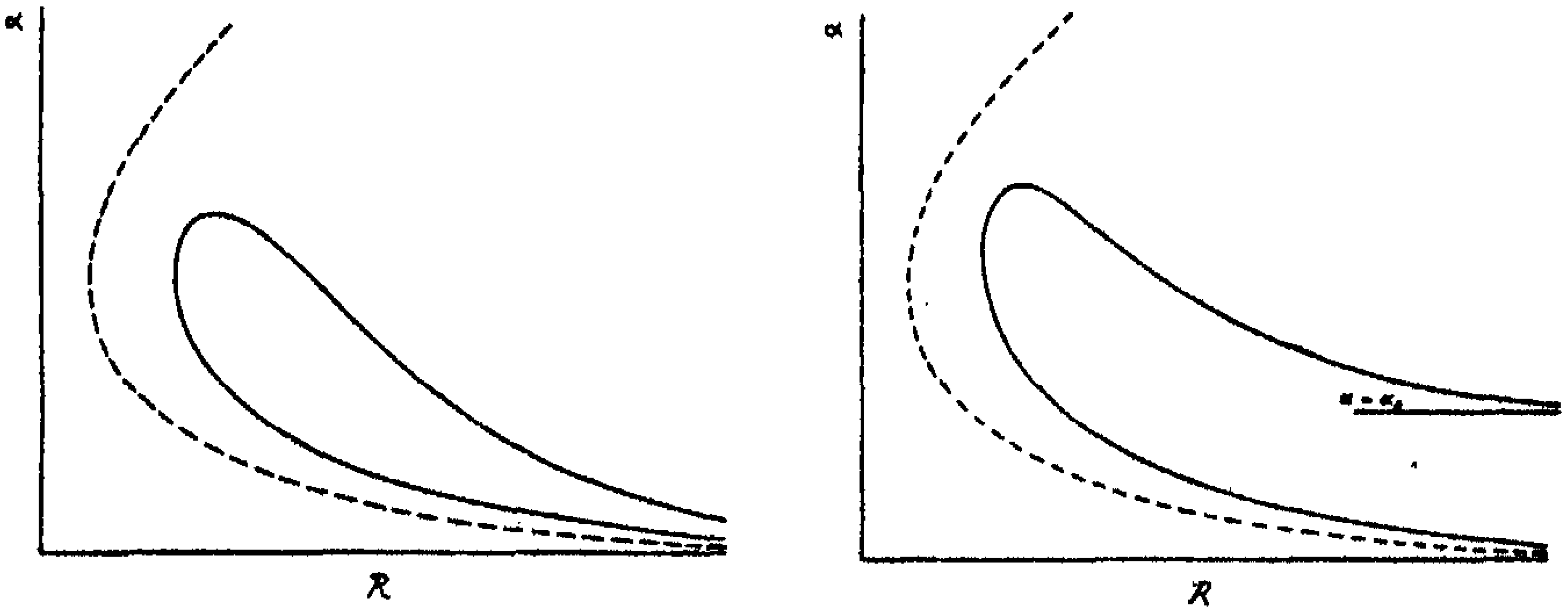


FIGURE 1 (a) FIGURE 1 (b)

General nature of the curve of neutral stability. The dotted curve is a curve of stability given by Synge.

Simple general formulae have also been derived to express the exact forms of those asymptotic branches in terms of the velocity curve. The fact that these two asymptotic branches join together to give a maximum α and a minimum R (instead of going to infinite α) can be inferred from a criterion of Synge.[6] It states that there is always stability ($c_i < 0$) if

$$(qR)^2 < (2\alpha^2 + 1)(4\alpha^4 + 1)/\alpha^2, \quad q = \max_{y_1 \leq y \leq y_2} |w'(y)|. \qquad (3.2)$$

Indeed, the minimum value of the Reynolds number on the neutral curve marks the beginning of instability and is therefore very important. Accordingly, the following approximate formulae have been derived for the evaluation of this minimum value:

$$R = \frac{30w_1'}{c^3}\left\{\frac{w_1'}{c}\int_{y_1}^{y_2} w^2 dy\right\}^{1/2}, \text{ for symmetrical profiles,} \tag{3.3}$$

$$R = \frac{25w_1'}{c^4}, \text{ for boundary-layer profiles,} \tag{3.4}$$

where w_1' is the value of $w'(y)$ at $w = 0$, and c is given by

$$-\pi w_1' \frac{w(y)w''(y)}{\{w'(y)\}^3} = 0.6 \text{ at } w(y) = c. \tag{3.5}$$

In these dimensionless expressions, the velocities are referred to the maximum velocity and the lengths to the distance from $w = 0$ to $w = 1$. Complete numerical calculation of the curve $c_i(\alpha, R) = 0$ has been carried out for the plane Poiseuille flow and the Blasius flow. The results are in general agreement with figure 1 (*a*). For the channel, the critical Reynolds number is found to be 16,000 as defined in terms of the maximum velocity and the width of the channel; for the boundary layer, it is found to be 400 as defined in terms of the free-stream velocity and the displacement thickness. The result in the latter case is in better agreement with Tollmien's original calculation[3] than with Schlichting's later one.[4] The maximum value of α is even larger than Tollmien's value, while Schlichting's value is smaller.

4. *Physical Interpretations and Concluding Discussions.*—We shall now try to understand the physical mechanism underlying the above results. Generally speaking, the investigation in an inviscid fluid brings out the rôle of pressure and inertial forces. The subsequent investigation of viscous forces will then settle the stability problem of a real fluid.

Equation (2.1) is the equation of vorticity. This suggests that the condition $w'' = 0$ in Rayleigh-Tollmien's criterion means an extremum of vorticity of the main flow $w(y)$. In fact, the physical mechanism can be understood by following this line of thought. Let us regard the field of flow as vorticity distribution in parallel layers, and imagine a disturbance consisting of the interchange of two fluid filaments in two different layers. These fluid elements will retain their original vorticities, and will therefore be present in layers with vorticities different from their own. It can be shown that these filaments will then be pushed back to their original layers if the vorticity gradient of the main flow has the same sign.[12] The flow is therefore stable. Indeed, if we consider an excess concentrated distribution of vorticity $\zeta(x, y)$ (a "vortex") upon a small region of a given field of vorticity, $\zeta_0(y)$, then the vortex is accelerated with an average acceleration

$$\bar{a} = \frac{1}{\Gamma}\int\int v^2(x, y)\zeta_0'(y)dxdy \tag{4.1}$$

in the direction of the positive y-axis, where Γ is the total strength of the vortex, $v(x, y)$ is the y-component of velocity due to it, and the integration is taken over the whole field of flow. A consideration of the sign of Γ and $\zeta_0'(y)$ will then substantiate the above statements. Thus, the stability in an inviscid fluid as governed by inertial and pressure forces can be discussed in terms of the gradient of vorticity of the main flow.

Another interpretation of the significance of the gradient of vorticity is to be found from Kelvin's "cat-eye" picture of a flow with neutral disturbance.[18] If we consider the vorticity of fluid elements along a closed stream line in that picture, we see that such a flow pattern is not possible without diffusion of vorticity by viscous forces, if the gradient of vorticity is finite at the critical layer where $w = c$.

This consideration immediately leads to a simple method of visualizing the extent of viscous forces in controlling the stability of fluid motion. If we consider the diffusion of vorticity from the critical layer for a period T of a neutral disturbance, the effective distance covered would be of the order of $(\nu T)^{1/2}$, ν being the kinematical viscosity. The ratio of this distance to that between the critical layer and the solid boundary can be easily shown to be

$$s = \sqrt{\frac{2\pi}{z^3}} \tag{4.2}$$

where $z = -\alpha_0\eta_1$ is directly related to the argument of the Hankel functions of Section 2, η_1 being the value of η at $y = y_1$. This parameter s has the value 0.7 at infinite Reynolds number on the lower branch of the neutral curves of figure 1. It decreases to about 0.5 at minimum Reynolds number and further decreases to zero as R becomes infinite along the upper branch. Referring to those figures, we note that for small values of s (small viscosity), the effect of viscosity is essentially destabilizing, since an increase of Reynolds number gives more stability (cf. Heisenberg's criterion of Section 3). For large values of s (large viscosity), the opposite is true. The physical mechanism can be described in the following way. It is known that the destabilizing effect is caused by phase shifts of the disturbance, which tends to build up its shear $-\rho\overline{u'v'}$ (u', v' being the components of velocity of the disturbance, and ρ the density of the fluid). This, in turn, transfers the energy of mean flow into that of small oscillations. The stabilizing effect is due to dissipation. Now, for small values of s, we may think of the flow as having two thin viscous layers, one at the wall, the other at the layer where $w = c$; the effect of dissipation is relatively unimportant, and the resultant influence of viscous forces is destabilizing. When s exceeds 0.5, these two layers may be regarded as having joined each other into one viscous region; the effect of dissipation becomes important, and the resultant influence is stabilizing.

The viscous forces upset the decisive nature of an extremum of vorticity for the determination of the stability of the flow of a real fluid, as one can see from a comparison of figure 1 (*a*) (no extremum) and figure 1 (*b*) (with extremum). Thus, there is no drastic change in the stability characteristics of a boundary layer as the pressure gradient changes from a slightly positive value to a slightly negative value (as one would expect if one tries to apply Rayleigh-Tollmien's criterion to a real fluid). However, the appearance of the point of inflection in a velocity curve does indicate an increase of instability. It can be shown, by using (3.3), (3.4) and (3.5), that the critical Reynolds number will be decreased. Also, for a given Reynolds number, the range of wave-lengths of unstable disturbances will be increased. For a boundary layer with adverse pressure gradient, a further indication of increasing instability is given by the decrease of w_1'. Indeed, as $w_1' \rightarrow 0$, the approximate formulae (3.4), (3.5) give $c \rightarrow 1$, $R \rightarrow 0$. This is, of course, an extrapolation of those formulae, but the conclusion stands in qualitative agreement with the usual notion that the flow would be violently unstable in such a case (separation).

5. *Transition to Turbulence.*—Finally, we want to touch briefly the question of transition to turbulence. A linearized theory of hydrodynamic stability is of course not quite adequate to account for such a phenomenon. But if we allow our theory to be extrapolated to the beginning of non-linear effects, and assume the subsequent development into turbulence to be very rapid, there is the possibility of determining approximately the Reynolds number of transition. For flow in boundary layers, it is interesting to note that the disturbance in the free stream has an isotropic nature in two dimensions. There may be therefore the possibility of introducing the results of the theory of instability in boundary layers into Taylor's theory of the transition Reynolds number,[5] which appears to be quite successful when the transition is caused by the turbulence in the free stream.

The author wishes to express his sincere gratitude to Professor Theodore von Kármán for suggesting this problem and for his invaluable help throughout this work; to Professor Clark B. Millikan for helpful suggestions and for using some of his unpublished notes, from which much inspiration and suggestion has been derived; to Professors P. S. Epstein, H. Bateman, Dr. Hans W. Liepmann and several of his friends for many valuable discussions and suggestions.

[*] The present article is a brief report of some results contained in the author's dissertation approved by the California Institute of Technology for the degree of Doctor of Philosophy.

[1] Heisenberg, W., *Ann. d. Phys.*, **74**, 577–627 (1924), especially p. 599.

[2] Tietjens, O., *Zeits. Angew. Math. Mech.*, **5**, 200–217 (1925).

[3] Tollmien, W., *Göttinger Nachrichten*, 21–44 (1929).

[4] Schlichting, H., *Ibid.*, 181–208 (1933).

[5] Taylor, G. I., *Proceedings of the Fifth International Congress for Applied Mechanics*, Cambridge, Mass., U.S.A. (1938), pp. 304–310, especially equation (45), p. 307.

[6] Synge, J. L., *Semicentennial Publications Amer. Math. Soc.*, Vol. II: Semicentennial Addresses, pp. 227–269 (1938); especially equation (11.23), p. 258.

[7] The conclusion that the plane Poiseuille flow may be unstable with respect to certain disturbances is in contradiction with that of Noether, F. (ZaMM, **6**, 232–243) (1926), and Pekeris, C. L. (*Jour. Aero. Sci.*, **5**, 237–240) (1938). The untenable points in their investigation are discussed in the present investigation.

[8] Squire, H. B., *Proc. Roy. Soc. London*, **A142**, 621–628 (1933).

[9] Rayleigh, Lord, *Scientific Papers*, Vol. 1, pp. 474–487.

[10] Tollmien, W., *Göttinger Nachrichten, Neue Folge*, **1**, 70–114 (1935).

[11] Rayleigh's treatment of this problem using broken linear velocity distributions has been denounced by Heisenberg, but his proof of the necessary condition is acceptable, except for the modification to be discussed below.

[12] The fact that an element with excessive vorticity is accelerated toward the region of higher vorticity was recognized and suggested to the author by Th. v. Kármán in connection with the theory of vorticity transfer in a fully developed turbulent flow.

[13] Kelvin, Lord, *Math. Phys. Papers*, Vol. 4, pp. 186–187.

FLAT SPACE-TIME AND GRAVITATION

By George D. Birkhoff

Department of Mathematics, Harvard University

Communicated September 7, 1944

If one admits that physical events take place in a 4-dimensional space-time continuum (an idea abandoned in current quantum-mechanical theory) there are three interesting possibilities: classical space and time; flat or electromagnetic space-time; curved space-time. The appropriate corresponding mathematical languages are, respectively, those of 3-vectors, 4-vectors and 4-tensors.

In a certain sense the flat space-time, characteristic of the so-called special theory of relativity, is just as *absolute* as classical space and time, since the coördinates t, x, y, z require exactly 10 arbitrary constants for their complete specification in both cases. But, in the framework of flat space-time, the fundamental electromagnetic equations of Maxwell and Lorentz lose the artificiality which they possess in classical space and time.

The initial attempts to incorporate gravitational phenomena in flat space-time were not satisfactory. Einstein turned to the curved space-time suggested by his principle of equivalence, and so constructed his general theory of relativity. The initial predictions, based on this celebrated theory of gravitation, were brilliantly confirmed. However, the theory has not led to any further applications and, because of its complicated mathematical character, seems to be essentially *unworkable*. Thus curved space-time has come to be regarded by many as an auxiliary construct (Larmor) rather than as a physical reality.

In my opinion, the failure of the early attempts by Nordström and others to develop a theory of gravitation in flat space-time is to be attributed to the fact that a fundamental theoretic requirement was overlooked, namely, that the disturbance velocity in matter must be that of light.

With this requirement in mind, I have recently been led to a very simple theory of gravitation in flat space-time, concordant with all known gravitational phenomena and free of arbitrary constants. This theory was presented first in very brief form in 1942 at Tonanzintla, Mexico, and has been developed further in a Note in these PROCEEDINGS.[1] Furthermore, attention is to be directed to a Note by A. Barajas,[2] called forth by a review of H. Weyl, and to an article by A. Barajas, C. Graef, M. Sandoval Vallarta and myself, taking up the new theory from the physical point of view.[3] A very significant application of the theory to the two-body problem by Graef will be published shortly.[4]

Unfortunately, the foundation of the theory has not so far been adequately presented in its philosophic, postulational and mathematical aspects. My colleague, Professor Barajas, and I are planning to publish an extensive article remedying this deficiency.

The aim of the present Note is to present briefly these foundational considerations as I see them. It is especially necessary to do so in order to avoid further misunderstanding of the new theory. For example, Weyl says very recently,[5] referring to my theory: "Their [i.e., 'the field equations'] most general static centrally symmetric solution involves 3 arbitrary constants a, b, l. . . . From the present standpoint this is a serious disadvantage of B." His assertion is wrong since the general exact solution for the gravitational potentials h_{ij} is

$$h_{ij} = \delta_{ij}\frac{m}{r}$$

where r stands for radial distance, δ_{ij} is the familiar Kronecker δ, and the mass m is the single arbitrary constant which enters. This exact solution plays in my theory a rôle analogous to that of the Schwarzschild solution in the theory of Einstein. Weyl has overlooked the salient fact that the central body is composed of the basic "perfect fluid."

The proposed theory of gravitation in flat space-time may be characterized as follows in its fundamental features:

1. In the 4-dimensional framework of flat space-time, matter is regarded as the occupant of certain tubular regions made up of the worldlines of

identifiable points. Point-particles are abandoned once and for all, except as a limiting possibility.

Thus there is a duality of matter and space-time in my theory, whereas the monistic concept of space-time conditioned by matter prevails in Einstein's theory. Whichever point of view is destined to figure in ultimate physical theory, it seems the part of obvious common sense to explore both possibilities fully.

The simplest available form of matter in flat space-time is that characterized by a certain stream 4-vector v^i of space type, i.e., with

$$\rho^2 = (v^1)^2 - (v^2)^2 - (v^3)^2 - (v^4)^2 > 0;$$

here ρ is the scalar length of the 4-vector v^i. Or, alternatively, we may think of matter of this kind as characterized by a density ρ and a velocity 4-vector u^i, where $v^i = \rho u^i$ and

$$(u^1)^2 - (u^2)^2 - (u^3)^2 - (u^4)^2 = 1.$$

The principle of local causation is taken to hold for an isolated portion of this kind of matter, as embodied in the differential equations:

$$\frac{\partial v^i}{\partial t} = F^i\left(v^1, \ldots v^4, \frac{\partial v^1}{\partial x}, \ldots \frac{\partial v^4}{\partial z}\right)$$

where $t = x^1, x = x^2, y = x^3, z = x^4$, and where F^i are taken to be rational and integral in the partial derivatives involved. These four differential equations assert that the time rates of change of the components of the stream vector are functions of these components and their space rates of change, being rational and integral in the latter.

The requirements of the underlying language of 4-vectors, based on the Lorentz group, then indicate as the *only* available possibility, when referred to instantaneous rest coördinates ($v^1 = \rho, v^2 = v^3 = v^4 = 0$):

$$\frac{\partial v^1}{\partial t} = F(\rho)\left(\frac{\partial v^2}{\partial x} + \frac{\partial v^3}{\partial y} + \frac{\partial v^4}{\partial z}\right),$$

$$\frac{\partial v^i}{\partial t} = G(\rho)\frac{\partial v^1}{\partial x^i} \qquad (i = 2, 3, 4).$$

These equations involve a pair of arbitrary functions $F(\rho)$, $G(\rho)$.

By appropriate normalization of the scalar ρ (that is by multiplicative modification of the 4-vector v^i) the preceding equations may be expressed in the normal form

$$\frac{\partial T^{i\alpha}}{\partial x^\alpha} = 0 \qquad (T^{ij} = v^i v^j - p(\rho)g^{ij})$$

where $g^{11} = -g^{22} = -g^{33} = -g^{44} = 1, g^{ij} = 0$ for $i \neq j$. Here there appears only one arbitrary function $p(\rho)$. This normalized density ρ is determined up to a unit of magnitude, while the scalar $p(\rho)$ is determined up to an additive constant.

2. It is granted that free equilibrium is possible for a certain equilibrium density, $\rho_0 > 0$, i.e., that a possible state is $v^i = v_0^i$ provided only that we have

$$(v_0^1)^2 - (v_0^2)^2 - (v_0^3)^2 - (v_0^4)^2 = \rho_0^2.$$

Because of the thoroughgoing analogy of the equations obtained with those of a homogeneous adiabatic fluid in classical physics, it is natural to assume that along a free boundary we have always $\rho = \rho_0$, and furthermore that at collision ρ takes on the same value on both sides of the common boundary until separation occurs for $\rho = \rho_0$ subsequently.

At this stage the behavior of a collection of freely moving portions of this "fluid" has been completely specified, whether or not collisions occur. The equations involved present a more familiar aspect if the velocities u^i are introduced, with

$$T^{ij} = \rho u^i u^j - p(\rho) g^{ij}.$$

This type of equations has always been taken to be appropriate for a general homogeneous adiabatic fluid in flat space-time. The symmetric tensor T^{ij} has been called the energy tensor; ρ, the density; and $p(\rho)$, the pressure.

3. Such a fluid has the property that a certain divergence vanishes:

$$\frac{\partial}{\partial x^\alpha}(e^{-\int dp/\rho}\, v^\alpha) = 0$$

This ensures that the 3-dimensional integral

$$\int \rho e^{-\int dp/\rho} dv$$

over the rest-volume is invariable.[6]

Keeping in mind the hydrodynamic analogy, it appears to be absolutely essential to suppose that this divergence vanishes under all conditions. Otherwise the fluid might undergo a full cyclic return to a set of initial velocities without a cyclic return of densities, such as is always observed to occur with ordinary matter. This requirement means that we have always

$$v_\beta \frac{\partial T^{\beta\alpha}}{\partial x^\alpha} \equiv 0.$$

4. There is, however, a fundamental theoretic difficulty in the theoretic employment of the general adiabatic fluid. In fact if two portions of the fluid collide with oppositely directed velocities nearly equal to that of light

(1)[7] the equations of motion will break down if the disturbance velocity of the fluid is less than that of light. Since it is physically inadmissible that this velocity v (relative to rest coördinates) exceed that of light, we are led to require that this velocity, namely,

$$v = \sqrt{dp/(d\rho - dp)},$$

equals that of light, so that at all densities $dp/d\rho = {}^1/_2$.

Thus we obtain by integration the only physically allowable constitutive equation $p = \rho/2$. The corresponding special form of the general adiabatic fluid is called the *perfect fluid*.

As far as this determination of T^{ij} is concerned, we might have written more generally $p = {}^1/_2\rho + c$, and so have obtained an additional term of the form $-cg^{ij}$ in T^{ij}. This modification would not affect the equations of motion, however. Reason will be given later on for the special choice made of $c = 0$ in fixing completely the energy tensor T^{ij}.

For the perfect fluid the invariable integral reduces to the simple form

$$\int \sqrt{\rho}dv$$

where dv is the 3-dimensional rest volume.

The perfect fluid may be looked upon as the counterpart in flat space-time of the homogeneous adiabatic incompressible fluid in classical space and time, which has infinite disturbance velocity. Physically speaking, the perfect fluid is very nearly incompressible and thus possesses very nearly invariable mass $\int \rho dv$.

If electricity of density σ be attached to the perfect fluid, the ratio $\sigma/\sqrt{\rho}$, called the substance coefficient, remains forever constant along any worldline.

In what follows the perfect fluid is regarded as the single primordial form of matter.

5. Suppose now that the perfect fluid, with energy tensor

$$T^{ij} = v^i v^j - {}^1/_2(v_\alpha v^\alpha)g^{ij} \qquad (1)$$

is not free but is subject to body forces. Formally, the suggestion is obvious that we *define* a force vector f^i by means of

$$\frac{\partial T^{i\alpha}}{\partial x^\alpha} = f^i \qquad (2)$$

where because of the postulated invariance of the integral

$$\int \sqrt{\rho}dv \qquad (3)$$

we have

$$f_\alpha v^\alpha \equiv 0 \qquad (4)$$

Thus the force vector is required to be identically orthogonal to the stream (or velocity) vector. It may be recalled that the acceleration vector a^i along any worldline has the same property.

In view of the identity (4) it is clear that f^i cannot be independent of the vector v^i. For the case of electrically charged matter f^i is known to be linear homogeneous in the 4-vector v^i and identically orthogonal to it. For the gravitational forces it is the simplest possible hypothesis to assume that in a purely gravitational field f^i is homogeneous and quadratic in v^i and proportional to ρ. We write therefore in that case

$$f^i = \rho\varphi^i_{\alpha\beta}v^\alpha v^\beta \tag{5}$$

where the components of the tensor $\varphi^i_{jk} = \varphi^i_{kj}$ are functions of t, x, y, z defined throughout space-time.

Without going into any detail it is to be stressed that this is the most natural assumption which can be made about the manner in which f^i depends on the stream vector v^i. This is especially the case in view of the reversibility of gravitational phenomena in time.

6. At this stage of our genetic account of the theory under consideration, in full accordance with the tradition of the past, it will be supposed to begin with that non-gravitational forces, such as electrical forces, need not be considered in gravitational problems. This hypothesis is justified by the more complete form of the theory given in the Note already alluded to.[1]

For the case of a rest system the x, y, z components of the gravitational forces take the form

$$f_x = \rho\varphi^2_{11}, \qquad f_y = \rho\varphi^3_{11}, \qquad f_z = \rho\varphi^4_{11} \tag{5'}$$

with $f_t = 0$ of course.

In the corresponding Newtonian situation the force components are:

$$f_x = \rho\frac{\partial g}{\partial x}, \qquad f_y = \rho\frac{\partial g}{\partial y}, \qquad f_z = \rho\frac{\partial g}{\partial z},$$

where g is the Newtonian potential defined by Poisson's equation

$$\frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2} + \frac{\partial^2 g}{\partial z^2} = -4\pi\rho, \text{ or } 0 \text{ in empty space},$$

together with the requirement that g is finite (or vanishes) at infinity.

By formal analogy one is led to require that φ^i_{jk} are linear homogeneous in the first derivatives of the gravitational potential defined by means of an appropriate Poisson equation.

A very simple possibility would be to set up a scalar gravitational potential h defined by

$$\frac{\partial^2 h}{\partial t^2} - \frac{\partial^2 h}{\partial x^2} - \frac{\partial^2 h}{\partial y^2} - \frac{\partial^2 h}{\partial z^2} = 8\pi T, \text{ or } 0 \text{ in empty space},$$

where we have written

$$T = g_{\alpha\beta} T^{\alpha\beta} = \rho/_2$$

for the contracted energy tensor. However, with this form of f^i it is not possible to fulfil the vital condition of orthogonality, embodied in (4).

The alternative, equally simple, hypothesis from the formal point of view is to take as the analog of the Poisson equation

$$\left[\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2}\right] h_{ij} = 8\pi T_{ij}, \text{ or } 0 \text{ in empty space,} \quad (6)$$

with the additional requirement that h_{ij} are finite (or vanish) at infinity. The gravitational potential h_{ij} thus introduced is a symmetric tensor of the second order. No essential limitation is introduced by use of the special numerical factor 8π on the right.

It is now readily seen that the same condition of orthogonality (4) leads to the following uniquely determined form for the gravitational force vector f^i:

$$f_i = \left(\frac{\partial h_{i\alpha}}{\partial x^\beta} - \frac{\partial h_{\alpha\beta}}{\partial x^i}\right) v^\alpha v^\beta \quad (7)$$

The theory has now been completely formulated for the limited form in which only gravitational forces enter. It may be extended so as to include electrical and atomic forces in a natural and interesting manner (see reference 1 and Section 11 below). In the limited but important form now under consideration, with use of the usual tensorial subscript and superscript notation, the theory is embodied in the pair of equations

$$\begin{aligned} \frac{\partial T^{i\alpha}}{\partial x^\alpha} &= g^{i\gamma}\left(\frac{\partial h_{\gamma\alpha}}{\partial x^\beta} - \frac{\partial h_{\alpha\beta}}{\partial x^\gamma}\right) v^\alpha v^\beta, \\ g^{\alpha\beta}\frac{\partial^2 h_{ij}}{\partial x^\alpha \partial x^\beta} &= 8\pi T_{ij}, \text{ or } 0 \text{ in empty space,} \end{aligned} \quad (8)$$

where T^{ij} is given by equation (1).

At this stage it is clear why it was natural to take the undetermined constant c in T^{ij} to be 0. Otherwise we should have had

$$T^{ij} = v^i v^j - {}^1/_2(v_\alpha v^\alpha) g^{ij} - c g^{ij} = \rho(u^i u^j - {}^1/_2 g^{ij}) - c g^{ij}$$

and we should obtain as the Poisson equation for the limiting case $\rho = 0$, inside of the fluid,

$$g^{\alpha\beta}\frac{\partial^2 h_{ij}}{\partial x^\alpha \partial x^\beta} = -8\pi c g_{ij},$$

which seems inappropriate unless c is 0.

On the basis of the theory thus obtained it is found that isolated bodies have a static, centrally symmetric distribution, and that a collection of spherical bodies move with very high approximation according to the Newtonian law of gravitation, relative to a convenient rest system.

7. Let us now turn to the case of a static, centrally symmetric distribution of the perfect fluid. It is not difficult to determine the radial distribution of the density ρ, but for our present purposes it suffices to observe that we have precisely (r, radial distance),

$$T^{ij} = {}^1/_2\rho(r)\delta_{ij}.$$

Our extended Poisson equation takes the form

$$\frac{\partial^2 h_{ij}}{\partial x^2} + \frac{\partial^2 h_{ij}}{\partial y^2} + \frac{\partial^2 h_{ij}}{\partial z^2} = -4\pi\rho(r)\delta_{ij}, \text{ or } 0 \text{ in empty space.}$$

Here, of course, the boundary of the (spherical) distribution occurs for some value r_0 of r where $\rho(r_0) = \rho_0$. Thus we find as the *exact* solution for the centrally symmetric case

$$h_{ij} = \frac{m}{r}\,\delta_{ij}$$

in empty space ($r > r_0$), involving, as the single arbitrary constant, the mass m of the fluid.

In this way we obtain the gravitational potentials around a static, centrally symmetric body like the Sun. These potentials are not observably affected by random atomic motion.

8. Now consider a comparatively small approximate sphere of the perfect fluid, forming in the limit a kind of ideal particle of mass 0. First, we have essentially $\rho = \rho_0$ throughout, so that we may write

$$\frac{\partial T^{i\alpha}}{\partial x^\alpha} = \rho_0 \frac{\partial u^i u^\alpha}{\partial x^\alpha} = \rho_0 u^\alpha \frac{\partial u^i}{\partial x^\alpha}$$

with high approximation, since (4) yields with similar approximation

$$\frac{\partial u^\alpha}{\partial x^\alpha} = 0.$$

But we have

$$u^\alpha \frac{\partial u^i}{\partial x^\alpha} = \frac{du^i}{ds} = \frac{d^2 x^i}{ds^2} \qquad (ds^2 = dt^2 - dx^2 - dy^2 - dz^2)$$

along the worldline of the ideal particle, so that we may write

$$\frac{\partial T^{2\alpha}}{\partial x^\alpha} = \rho_0 \frac{d^2x}{ds^2}, \qquad \frac{\partial T^{3\alpha}}{\partial x^\alpha} = \rho_0 \frac{d^2y}{ds^2}, \qquad \frac{\partial T^{4\alpha}}{\partial x^\alpha} = \rho_0 \frac{d^2z}{ds^2}.$$

Thus there are obtained the differential equations of motion for a comparatively small body which moves in the field of a larger central body. These lead to essentially the same result for the perihelial advance of a planet and for the bending of a light photon in the field of the Sun as does the theory of Einstein.

Furthermore, if we assume that the Planck formula

$$E = h\nu$$

determines the frequency of radiation ν, the result of Einstein for the "red shift" in light reaching the Earth from the Sun is obtained. Inasmuch as the precise mechanism of radiation is unknown, it seems more correct to employ this basic formula, than to give an explanation in which the light-carrying photon plays no rôle, such as is afforded by the Einstein theory.

9. A real test of the availability of the new theory in other directions is afforded by the problem of two or more bodies.

As a first approximation to this problem, it is natural to consider the limiting case of n "ideal particles" of masses $m_1, m_2 \ldots m_n$, respectively, obtained by taking the equilibrium density ρ_0 to be very large. It is clear that in the neighborhood of each particle P_i of mass m_i, the corresponding gravitational potentials should have a principal part

$$\frac{m_i}{r_i}\delta_{ij}$$

in instantaneous rest coördinates.

Graef has already shown (see references 3 and 4) that the calculations involved can be effectively carried out in the case of two bodies of masses m_1 and m_2. Presumably his method can be extended to the case of more than two bodies. Furthermore it should be possible to investigate in a similar spirit all of the fine-structure corrections to the Newtonian theory which lie within the limits of observation.

10. In order to generalize the theory so as to admit cosmological terms one has only to write the Poisson equation in the form

$$\left[\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2}\right]h_{ij} = 8\pi T_{ij} + Kg_{ij}$$

where K is the small cosmological constant. This really means that we allow a form of energy tensor T^{ij} containing the term $-cg^{ij}$ previously indicated, with $c = -K/8\pi$ small but not 0.

The conditions at ∞ have then clearly to be lightened to the form that

h_{ij} becomes regularly infinite to the second order at infinity, in which case it is concluded that for a uniquely determined space-time origin we have

$$h_{ij} = h^*_{ij} + \frac{K}{8}(t^2 - x^2 - y^2 - z^2)g_{ij}$$

where h^*_{ij} satisfy the previously indicated form of the Poisson equation and boundary conditions. In this case there is obtained an expanding universe, about the space-time origin $t = x = y = z = 0$ in the flat space-time under consideration.

11. In the general case there is an electromagnetic 4-potential φ_i satisfying Maxwell's equations, and an atomic potential ψ constant along every worldline. This leads to the complete theory with (covariant) force components:

$$f_i = \rho\frac{\partial\psi}{\partial x^i} + \sigma\left(\frac{\partial\varphi_i}{\partial x^\alpha} - \frac{\partial\varphi_\alpha}{\partial x^i}\right)u^\alpha + \rho\left(\frac{\partial h_{i\alpha}}{\partial x^\beta} - \frac{\partial h_{\alpha\beta}}{\partial x^i}\right)u^\alpha u^\beta \qquad (9)$$

where the terms corresponding to atomic, electric and gravitational forces are homogeneous of degrees 0, 1 and 2, respectively, in the velocity vector, and are linear homogeneous in the first partial derivatives of the corresponding potentials ψ, φ^i and h^{ij}, respectively.

It will require further mathematical investigation in order to determine the serviceability of this conjectural theory in the domain of atomic physics. I have previously indicated how an equation much like the Schrödinger wave equation may be obtained on the basis of the atomic potential ψ.[8] However, my attempt was based on a background of curved space-time, and I had not then realized that the potentials must all be of zero dimensions, so that I used $\partial\psi/\partial x^i$ in place of $\rho\partial\psi/\partial x^i$.

12. Professors Barajas, Graef, Sandoval Vallarta and I are examining further these and other physical problems in the light of the new theory. Meanwhile, it seems clear that the theory promises well and deserves careful study because of its striking simplicity, completeness and mathematical consistency. Furthermore, as will be seen from what precedes, the theory is independent of all ideas of curved space-time and of the corresponding Einstein theory.

No doubt, in view of the substantial successes of the Einstein theory, it is worth while to attempt to reflect that theory on to flat space-time, and so to obtain a degenerate theory, which in a certain sense is only the shadow of a shadow. However, the objections made by Barajas[2] to the form of degenerate theory considered by Weyl[5] seem to be substantially valid.[9]

As far as I can see, the Einstein principle of equivalence ("that inertia and gravitation are one," Weyl, loc. cit.) is at bottom a mathematical principle signifying only that certain equations in the Einstein theory are linear homogeneous in the density of matter—a fact just as true of the theory of gravitation here proposed as of the general theory of relativity. To me there is only the following mathematical fact in the comparison of the basic points of view of the Einstein theory and my own: in his theory there is no underlying framework of independent variables t, x, y, z valid throughout space-time, such as are present in the theory based on flat space-time. The real question is whether or not the theory based on such special coördinates is simpler and more useful for the description and prediction of the physical facts. This is not a question to be decided by *a priori* considerations. What is required is rather a study of the new theory and its physical applications.

---

[1] Birkhoff, G. D., "Matter, Electricity and Gravitation in Flat Space-Time," these PROCEEDINGS, **29**, 231 (1943). My lecture at Tonanzintla is about to appear under the title "El Concepto de Tiempo y la Gravitacion," *Boletín de la Sociedad Matemática Méxicana*, **1**, No. 4 (1944).

[2] Barajas, A., "Birkhoff's Theory of Gravitation and Einstein's for Weak Fields," these PROCEEDINGS, **30**, 54 (1944).

[3] Barajas, A., Birkhoff, G. D., Graef, C., and Sandoval Vallarta, M., "On Birkhoff's New Theory of Gravitation," *Physical Review*, **66**, 138 (1944).

[4] In vol. 1, No. 5 (1944) of the *Boletín de la Sociedad Matemática Méxicana*.

[5] Weyl, H., "Comparison of a Degenerate Form of Einstein's with Birkhoff's Theory of Gravitation," these PROCEEDINGS, **30**, 205 (1944).

[6] See for instance, Birkhoff, G. D., *Relativity and Modern Physics*, Chaps. VII and XI, Cambridge, 1923, 1927.

[7] The velocity of light c is 1 since the lightsecond is regarded as the unit of distance.

[8] Cf. two notes in these PROCEEDINGS, **13**, 160, 165 (1927).

[9] In his Note Barajas is more than fair to the "degenerate theory", which, strictly speaking, is no more usable than is the early Nordström theory. See M. Wyman, *Math. Rev.* **5**, 218 (1944). Since *all* of the relativistic theories of gravitation take the classical Newtonian theory as prototype, the formal resemblance between them is inevitably considerable. This fact is stressed, for example, in my article, *Newtonian and Other Forms of Gravitational Theory*, Scientific Monthly, **58**, 49 and 136 (1944). It is to be looked upon as the source of the formal resemblance between Einstein's general theory of relativity, based on curved space-time, and my own theory, based on flat space-time, which Weyl refers to.

PROCEEDINGS
OF THE
NATIONAL ACADEMY OF SCIENCES

Volume 30 November 15, 1944 Number 11

Copyright 1944 by the National Academy of Sciences

EXPERIMENTS ON SEXUAL ISOLATION IN DROSOPHILA. III. GEOGRAPHIC STRAINS OF DROSOPHILA STURTEVANTI

BY TH. DOBZHANSKY

DEPARTMENT OF ZOÖLOGY, COLUMBIA UNIVERSITY

Communicated September 28, 1944

When females of different geographic strains of the same species are confined with males of one of these strains, the matings may occur either at random or preferentially, one kind of females being inseminated more frequently than the other. Relatively slight mating preferences have been observed in Brazilian and Guatemalan *Drosophila willistoni*,[1] and some quite striking preferences have been found in Brazilian, Guatemalan and Mexican strains of *D. prosaltans*.[2] Most of these intra-specific mating preferences are one-sided: greater proportions of females of one than of the other strain are inseminated by males of both strains. Data reported in the present article show, however, that two-sided preferences for homogamic matings, which are, in general, characteristic of interspecific crosses, occur also within the species *D. sturtevanti* Duda (= *D. biopaca* Sturtevant).

The strains of *Drosophila sturtevanti* came from Tamazunchale, state of San Luis Potosi, Mexico (obtained through the courtesy of Professor J. T. Patterson), Quiriguá, Guatemala (obtained through the courtesy of Professor A. H. Sturtevant), Belem, state of Pará, Rio de Janeiro, Federal District, and Bertioga, state of São Paulo, Brazil. The flies from all these strains are morphologically similar, intercross readily and give rise to fertile F_1 and F_2 progenies. The technique used in the experiments with *D. sturtevanti* is identical with that used for *D. prosaltans*,[2] except that *D. sturtevanti* females are dissected on the sixth or the seventh day after the making of the crosses. Sexual activity in this species starts late, in fact later than in other species of Drosophila with which the writer is familiar. Before the sixth day scarcely any females are inseminated, but on the eighth day only few remain not fertilized. The tubular (ventral) receptacle in *D. sturtevanti* is an extremely long tube, much longer than the body, resembling a skein of wool that is twisted, as a whole, into paired spirals lying on both sides of the vagina. Live spermatozoa can be seen in some, but usually not in all, coils. The large but transparent chitinous spermathecae are, in inseminated females, only partly filled with a tangled mass of sperm.

The data are summarized in table 1. In every one of the 20 crosses a positive isolation index has been obtained, which suggests that males of all strains inseminate more of their own females than of those of other strains. The statistical significance of the observed deviations from random mating is measured by the χ^2's; values higher than 6.6 (corresponding to the 0.01 probability level) have been observed in 10 crosses, and values smaller than 2.7 (corresponding to the 0.10 probability level) in only 5 crosses. The highest isolation indices are found in the crosses of strains from the two geographically most remote localities (Tamazunchale in Mexico and Bertioga in Brazil), but localities nearly as remote (Tamazunchale and Rio de Janeiro) fail to show statistically significant isolation indices. Strains from geographically relatively close localities may either interbreed at random (Rio de Janeiro and Bertioga) or may display a considerable isolation (Tamazunchale and Quiriguá). Strains which show no isolation from each other (Rio de Janeiro and Bertioga) may behave very differently with respect to other strains (Tamazunchale, Belem, or Quiriguá).

TABLE 1

NUMBERS OF FEMALES DISSECTED (n) AND PER CENT CARRYING SPERM (%) IN VARIOUS CROSSES OF *Drosophila sturtevanti*

FEMALES	MALES	HOMOGAMIC n	HOMOGAMIC %	HETEROGAMIC n	HETEROGAMIC %	χ^2	ISOLATION INDEX
Tamazunchale, Quiriguá	Tamazunchale	194	60.3	192	37.0	21.0	0.24
Tamazunchale, Quiriguá	Quiriguá	104	46.2	107	24.3	11.1	0.31
Tamazunchale, Belem	Tamazunchale	111	54.1	109	32.1	10.7	0.25
Tamazunchale, Belem	Belem	86	65.1	82	41.5	9.5	0.22
Tamazunchale, Rio	Tamazunchale	104	52.9	112	46.4	0.9	0.06
Tamazunchale, Rio	Rio	124	61.3	130	49.2	3.7	0.11
Tamazunchale, Bertioga	Tamazunchale	115	73.9	104	35.6	32.6	0.35
Tamazunchale, Bertioga	Bertioga	108	58.3	108	24.1	26.2	0.42
Quiriguá, Belem	Quiriguá	106	48.1	97	25.8	10.9	0.30
Quiriguá, Belem	Belem	94	54.3	99	24.2	18.3	0.38
Quiriguá, Rio	Quiriguá	99	56.6	96	43.8	3.2	0.13
Quiriguá, Rio	Rio	116	50.0	113	35.4	5.0	0.17
Quiriguá, Bertioga	Quiriguá	99	49.5	103	32.0	6.4	0.21
Quiriguá, Bertioga	Bertioga	106	59.4	109	26.6	23.7	0.38
Belem, Rio	Belem	160	55.6	157	49.0	1.4	0.06
Belem, Rio	Rio	98	65.3	106	39.6	14.5	0.24
Belem, Bertioga	Belem	101	50.5	100	39.0	2.7	0.13
Belem, Bertioga	Bertioga	91	41.8	88	31.8	2.0	0.14
Rio, Bertioga	Rio	106	51.9	106	50.0	0.1	0.02
Rio, Bertioga	Bertioga	101	50.5	106	48.1	0.1	0.02

The situation found in *D. sturtevanti* is different from those in *D. willistoni*[1] and *D. prosaltans*.[2] In *D. sturtevanti* the rule seems to be that strains from different regions show incipient sexual isolation—preference for homogamic matings and aversion from heterogamic ones. Random mating is the rule in *D. willistoni,* but both Guatemala and Brazilian males have in some crosses inseminated preferentially Brazilian females. In *D. prosaltans* there is a "hierarchy" of strains correlated with their geographic origin. Both northern and southern males inseminate preferentially northern females; therefore, in some crosses homogamic matings are more common than heterogamic ones (positive isolation indices), and in other crosses heterogamic matings are more frequent (negative isolation indices). The strains from Guatemala and Brazil show, however, a mutual aversion to interbreeding. *D. nebulosa* Sturtevant is another species in which preferences for homogamic as well as for heterogamic matings are observed. Strains of this species coming from state of San Luis Potosi, Mexico (kindly furnished by Professor J. T. Patterson), and from Belem do Pará, Brazil, have been used. The experiments on *D. nebulosa* were arranged exactly like those on *D. prosaltans* and *D. sturtevanti;* the dissections have been made on the fourth day after the crosses. A summary of the results is given in table 2. It is evident that when Brazilian and Mexican females are kept with Brazilian males, Brazilian females are inseminated preferentially; but when Mexican males are used there seems to be also an excess of inseminations of Brazilian females (a negative isolation index which is not quite significant statistically).

TABLE 2

NUMBERS OF FEMALES DISSECTED (*n*) AND PER CENT CARRYING SPERM (%) IN CROSSES OF *Drosophila nebulosa*

FEMALES	MALES	HOMOGAMIC *n*	HOMOGAMIC %	HETEROGAMIC *n*	HETEROGAMIC %	χ^2	ISOLATION INDEX
Mexican, Brazilian	Brazilian	126	78.6	133	41.4	37.1	0.31
Mexican, Brazilian	Mexican	113	42.5	113	55.8	4.0	−0.14

Two types of preferential mating have been observed in Drosophila. They are sexual isolation proper and one-sided mating preferences. Strains A and B are said to be sexually isolated if males of A mate chiefly with A, and males of B chiefly with B females (positive isolation indices). If the mating preference is one-sided, A males mate more frequently with A than with B females (positive isolation index), but B males show either no preference (isolation index close to zero) or a preference for A females (negative isolation index). Situations intermediate between true sexual isolation and one-sided mating preference occur. Theoretically, two strains may also show consistent preferences for heterogamic matings (negative isolation indices), but this has never been found in experiments.

Sexual isolation reduces the rate of gene exchange between the populations concerned; when the isolation is complete the gene exchange is altogether excluded. Sexual isolation is found, as a rule, in interspecific crosses. Rudiments of sexual isolation have been found also within a species, e.g., in crosses between certain strains of *D. miranda*,[8] *D. sturtevanti* and *D. prosaltans* from Brazil and Guatemala.[3] One-sided mating preferences are known in most geographic races of *D. prosaltans*, in *D. Willistoni*[1] and in *D. nebulosa*. Such mating preferences do not render the gene exchange impossible, because at least one of the two populations would continue to receive an inflow of genes from the other.

Since hybrids between two forms, each of which possesses a well-balanced genotype, may be ill balanced or unfit for survival, reduction or cessation of the gene exchange between populations may be adaptively important. Genetic variants which reduce the rate of the gene exchange may be favored by natural selection in populations exposed to hybridization. Reproductive isolation would be strengthened by natural selection until complete suspension of the gene exchange is attained. The development of all forms of reproductive isolation, including sexual isolation, between populations which have diverged sufficiently to make compromise genotypes unfavorable may be initiated and furthered by natural selection.[4, 5] On the other hand, Muller[6, 7] has pointed out that impediments to gene exchange may arise accidentally, as concomitants of the process of divergence of genetic systems.

These two suggestions may be complementary rather than conflicting. It is well known that natural selection operates with the genetic variants present in populations, but cannot create such variants. The variants are bound to arise, in a sense, accidentally, through mutation and recombination. The one-sided mating preferences of the kind observed in *D. prosaltans* would not, even if the geographic races involved were exposed to hybridization, constitute a barrier to gene exchange. Such mating preferences may, however, serve as raw materials from which true sexual isolation could develop by addition of genetic variants making the aversion to heterogamic matings mutual instead of unidirectional. Incipient sexual isolation is, indeed, found between strains of *D. prosaltans* from Brazil and Guatemala. Incipient sexual isolation is widespread among geographic strains of *D. sturtevanti*, but it is still neither strong enough to be a serious obstacle to gene exchange, nor ordered into a balanced system of isolation correlated with geography. One-sided mating preferences are very likely by-products of physiological racial differences, not necessarily related to the mating reactions. True sexual isolation develops if one-sided mating preferences are added up and coördinated by natural selection.

The great diversity and complexity of structures and habits connected with the search for mates, nest building, courtship and copulation in many groups of animals is very impressive. These structures and habits are often not only surprisingly complex but also constant within and characteristically different between species. Their origin in evolution seems difficult to explain. Their orderly complexity makes accidental origin extremely improbable. Natural selection would seem to be incapable of building them because of the apparent lack of adaptive value in most of these structures and habits. Darwin's theory of sexual selection was a partially successful attempt to resolve this difficulty, but the competition for mating postulated by the theory is relatively seldom observed. It is, therefore, desirable to keep in mind the possibility that species specific courtship and mating habits and associated structures may serve as "recognition marks" of the specific identity.[8] If so, these habits and structures may possess adaptive value as integral parts of the mechanism of sexual isolation, and for this reason may be developed by a process of selection. Such a modification of the theory of sexual selection would, it seems, avoid most of the difficulties confronted by its original, Darwinian, version.

It should also be kept in mind that the problem of sexual structures and habits which seem to be devoid of adaptive significance is only a part of the broader problem of the origin in evolution of the so-called "neutral" characters in general. It is customary to suppose that neutral characters become established in evolution because of being physiologically correlated with some adaptively important, although less readily perceptible, characters. Unfortunately, the postulated correlations can only seldom be demonstrated to exist. For some neutral characters a reasonable alternative would be to suppose that they function as recognition marks of the species during the courtship and mating processes, and, are, therefore, susceptible of modification and development by selection.

The assistance of Miss Irene Markreich in conducting the experiments is gratefully acknowledged.

Summary.—Geographic strains of *D. sturtevanti* show incipient sexual isolation. The extent of this isolation is, however, variable and is not strictly correlated with the geographic origin of the strains; some strains coming from remote localities show less aversion to interbreeding than strains from less remote localities. The origin in evolution and the biological functions of the sexual isolation proper, as well as of other forms of preferential mating, are discussed.

[1] Dobzhansky, Th., and Mayr, E., these PROCEEDINGS, **30,** 238–244 (1944).

[2] Dobzhansky, Th., and Streisinger, G., *Ibid.*, an article appearing in this issue.

[3] Dobzhansky, Th., and Koller, P., *Biol. Zentr.*, **58,** 589–607 (1938).

[4] Dobzhansky, Th., *Amer. Natur.*, **74,** 312–321 (1940).

[5] Dobzhansky, Th., *Genetics and the Origin of Species*, New York, 1941.

[6] Muller, H. J., *Biol. Reviews*, **14,** 261–280 (1939).

[7] Muller, H. J., *Biol. Symposia*, **6,** 71–125 (1942).

[8] Mayr, E., *Systematics and the Origin of Species*, New York, 1942.

EXPERIMENTS ON SEXUAL ISOLATION IN DROSOPHILA. II. GEOGRAPHIC STRAINS OF DROSOPHILA PROSALTANS

By Th. Dobzhansky and G. Streisinger

Department of Zoölogy, Columbia University, and The Biological Laboratory of the Long Island Biological Association, Cold Spring Harbor, N. Y.

Communicated September 28, 1944

It has been shown[1] that when males of Brazilian strains of *Drosophila willistoni* are confined with females of their own and of other Brazilian strains, the females are inseminated at random. But when Brazilian or Guatemalan males are kept with mixtures of Brazilian and Guatemalan females, there is a tendency for a greater proportion of the Brazilian than of the Guatemalan females to be inseminated. A more complex system of sexual preferences, which exists within the species *D. prosaltans* Duda (= *D. sellata* Sturtevant), is described in the present article.

Strains of *D. prosaltans* coming from the following localities (Fig. 1) have been used: Huichihuayan, state of San Luis Potosi, Mexico; Chilpancingo and Zopilote Canyon, state of Guerrero, Mexico; Guatemala City, Guatemala; Belem, state of Pará, Brazil; Bertioga and Iporanga, state of São Paulo, Brazil. The strains from Mexico and Guatemala were obtained through the courtesy of Professors J. T. Patterson and A. H. Sturtevant. The Brazilian strains were collected by the senior author. All seven strains intercross readily, produce fertile hybrids, and doubtless belong to the same species. Yet the peculiar chromosome structure found in the South-Brazilian strains (Bertioga and Iporanga)[2] is absent in the strains from Belem, Guatemala and Mexico, which have V-shaped X's and Y's, one pair of V-shaped autosomes, and one pair of rod-like autosomes. Furthermore, in the Mexican strains the basal bands on the abdominal tergites are broader, lighter in color and sharper in outline than those of the Brazilian ones;[3] the Guatemala strain is intermediate in this respect and closer to those from Brazil than to those from Mexico. Finally, under the conditions of the experiments, sexual activity in the Mexican strains starts on the second day and in the Brazilian strains on the third day after hatching from the pupae.

A slight modification of the experimental techniques used for *D. willistoni*[1] was followed in the work with *D. prosaltans*. Ten freshly hatched females of each of two strains—i.e., twenty females in all—are placed in a vial with ten freshly hatched males of one of these strains. The right wings are clipped in one class of females and the left wings in the other, or else in one class a wing is cut and in the other left intact. The vials are kept at $24^1/_2$°C. for about 3 days, where cultures with Mexican males are concerned, or for 4 days if Brazilian males are used. The females are then dissected, and their seminal receptacles are examined for sperm. The vials in which too few or too many (less than 20% or more than 80%) of the females are inseminated are discarded (such vials are infrequent). The tubular (ventral) receptacle in *D. prosaltans* is a very long tube, resembling a skein of wool that is folded, as a whole, into the shape of an M. Live spermatozoa lying parallel to the axis of the tube are easily seen. The spherical chitinous spermathecae contain a tangled mass of sperm in which only some moving threads can be seen.[4] If a female is killed while in coitu or shortly thereafter, the spermathecae and all but the proximal gyres of the tubular receptacle are free of sperm, but a viscous ejaculate is found in the vagina.

A summary of the data is presented in table 1. Numbers of the females dissected and percentages of them found inseminated are given for the homogamic (i.e., inseminated by males of the same strain) and the heterogamic (i.e., inseminated by males of the foreign strain) combinations. The χ^2's for the observed differences between the frequencies of the homogamic and the heterogamic matings, and the isolation indices[5] are also included. It is evident at a glance that preferential mating is the rule rather than an exception when different geographic strains of *D. prosaltans* are brought together, and that strong preferences for homogamic matings (positive isolation indices) as well as for heterogamic matings (negative isolation indices) are common. Taking the 0.05 or 0.01 probability levels (corresponding to χ^2's of 3.84 and 6.64) as the limits of significance, we find that 35 or 34 out of the 42 crosses made give results deviating from randomness of mating. Significant positive isolation indices are found in 23 or 24, and significant negative ones in 11 crosses.

The relationships observed can be described most simply as follows. The geographic strains examined form a hierarchic series: Chilpancingo $\geqslant$ Zopilote $>$ Huichihuayan $>$ Guatemala $>$ Belem $>$ Iporanga $\geqslant$ Bertioga. If males of a strain higher up in this series are confined with females of the same strain and of another strain lower down in the series, the frequency of homogamic matings exceeds that of the heterogamic ones, making the isolation index positive. But if males of a "low" strain are kept with a mixture of females of their own and of a "higher" strain, the frequency of heterogamic matings exceeds that of the homogamic ones, and the isolation index is negative. This regularity was found in 36 out of 42 crosses; the exceptions are discussed below.

Chilpancingo males always inseminate more of their own than of foreign females. This preference[6] is very strong in mixtures of Chilpancingo with Bertioga and Iporanga females, a little less strong in mixtures with Belem, Guatemala and Huichihuayan, and least strong in mixtures with Zopilote females. It may seem surprising, however, that, with the exception of the Zopilote males which show no preference, males of all other strains in-

TABLE 1

NUMBERS OF FEMALES DISSECTED (n) AND PER CENT CARRYING SPERM (%) IN VARIOUS CROSSES OF *Drosophila prosaltans*

FEMALES	MALES	HOMOGAMIC n	HOMOGAMIC %	HETEROGAMIC n	HETEROGAMIC %	χ^2	ISOLATION INDEX
Chilpancingo, Zopilote	Chilpancingo	60	88.3	65	60.0	12.8	0.19
Chilpancingo, Zopilote	Zopilote	91	57.1	90	56.7	0.0	0.00
Chilpancingo, Huichihuayan	Chilpancingo	72	95.8	70	32.9	61.7	0.49
Chilpancingo, Huichihuayan	Huichihuayan	48	31.2	47	87.2	30.8	−0.47
Chilpancingo, Guatemala	Chilpancingo	60	96.7	63	28.6	60.7	0.54
Chilpancingo, Guatemala	Guatemala	93	40.9	106	71.7	19.3	−0.27
Chilpancingo, Belem	Chilpancingo	91	89.0	86	26.7	70.5	0.54
Chilpancingo, Belem	Belem	111	27.0	102	74.5	47.9	−0.47
Chilpancingo, Bertioga	Chilpancingo	48	91.7	48	6.2	70.0	0.87
Chilpancingo, Bertioga	Bertioga	103	17.5	111	65.8	51.0	−0.58
Chilpancingo, Iporanga	Chilpancingo	57	100.0	57	14.0	85.9	0.75
Chilpancingo, Iporanga	Iporanga	85	35.3	89	67.4	18.0	−0.31
Zopilote, Huichihuayan	Zopilote	71	71.8	72	34.7	19.7	0.35
Zopilote, Huichihuayan	Huichihuayan	105	26.7	114	69.3	39.7	−0.44
Zopilote, Guatemala	Zopilote	65	75.4	59	18.6	40.1	0.60
Zopilote, Guatemala	Guatemala	67	28.4	75	58.7	13.0	−0.35
Zopilote, Belem	Zopilote	46	95.6	42	26.2	45.5	0.57
Zopilote, Belem	Belem	112	49.1	113	61.1	3.2	−0.11
Zopilote, Bertioga	Zopilote	77	72.7	76	19.7	43.3	0.57
Zopilote, Bertioga	Bertioga	105	32.4	108	67.6	26.4	−0.35
Zopilote, Iporanga	Zopilote	50	92.0	53	32.1	38.8	0.48
Zopilote, Iporanga	Iporanga	82	26.8	80	70.0	30.1	−0.45
Huichihuayan, Guatemala	Huichihuayan	92	71.7	92	38.0	21.1	0.31
Huichihuayan, Guatemala	Guatemala	79	36.7	80	48.7	2.6	−0.14
Huichihuayan, Belem	Huichihuayan	86	77.9	82	13.4	70.1	0.71
Huichihuayan, Belem	Belem	106	57.5	100	30.0	15.8	−0.31
Huichihuayan, Bertioga	Huichihuayan	74	58.1	74	17.6	25.8	0.54
Huichihuayan, Bertioga	Bertioga	100	32.0	96	51.0	7.1	−0.23
Huichihuayan, Iporanga	Huichihuayan	76	81.6	75	48.0	18.7	0.26
Huichihuayan, Iporanga	Iporanga	100	33.0	95	45.3	3.1	−0.16
Guatemala, Belem	Guatemala	103	55.3	102	10.8	46.5	0.67
Guatemala, Belem	Belem	99	55.6	106	16.0	35.0	0.55
Guatemala, Bertioga	Guatemala	51	64.7	51	11.8	30.2	0.69
Guatemala, Bertioga	Bertioga	78	66.7	80	41.3	10.2	0.24
Guatemala, Iporanga	Guatemala	50	60.0	52	11.5	26.2	0.68
Guatemala, Iporanga	Iporanga	164	60.4	155	38.7	14.9	0.22
Belem, Bertioga	Belem	122	53.3	116	36.2	7.0	0.19
Belem, Bertioga	Bertioga	91	41.8	94	21.3	9.1	0.32
Belem, Iporanga	Belem	94	45.7	94	50.0	0.3	−0.04
Belem, Iporanga	Iporanga	85	43.5	80	41.2	0.1	0.03
Bertioga, Iporanga	Bertioga	70	42.8	67	38.8	0.2	0.05
Bertioga, Iporanga	Iporanga	99	61.6	108	48.1	3.8	0.12

seminate Chilpancingo females in preference to females of their own strains. The preference is exhibited most strongly by males from Bertioga, which give a negative isolation index as low as -0.58.

Zopilote flies are about equivalent to Chilpancingo, but the positive as well as the negative isolation indices tend to be lower where Zopilote flies are concerned. Huichihuayan males prefer their own females to those of Guatemala, Belem, Bertioga and probably Iporanga, but inseminate Chilpancingo and Zopilote females in preference to their own. Hiuchihuayan females placed with Chilpancingo and Zopilote are discriminated against by males of the latter strains, but are preferred by Guatemala, Belem, Iporanga and Bertioga males in the presence of their own females.

Guatemala and Belem strains behave somewhat differently from those discussed above. As might be expected from the position of the Guatemala strain in the hierarchic series, Guatemala males inseminate their own females preferentially in mixtures with Belem, Bertioga and Iporanga, but prefer foreign females in mixtures with Chilpancingo, Zopilote and probably Huichihuayan. But Guatemala females are not inseminated by males of any strain in preference to females of their own strain. Belem, Bertioga and Iporanga males, which stand lower in the hierarchic series than Guatemala, inseminate a greater proportion of their own than of Guatemala females; Chilpancingo, Zopilote and Huichihuayan males do likewise. There is, consequently, a true sexual isolation—that is, a reciprocal aversion—between the Guatemala and the Brazilian strains, and if flies from Guatemala and Brazil were mixed the matings would be mostly homogamic. As shown above, there is no reciprocal aversion between Guatemala and Mexican strains, or between Mexican and Brazilian strains, Mexican females being always preferred. The Belem strain seems to behave in a manner similar to the Guatemala. In accordance with the position of this strain in the hierarchy, Belem males give negative isolation indices in mixtures with Mexican and their own females, but positive isolation indices in mixtures with Bertioga females. However, Bertioga males placed with Bertioga and Belem females inseminate the former preferentially. Matings occur more or less at random in mixtures of Belem and Iporanga, or of Bertioga and Iporanga flies.

The hierarchy of sexual preferences of strains of *Drosophila prosaltans* is clearly correlated with the geographic origin of these strains (Fig. 1). The three highest strains all come from Mexico, which is the northern limit of the distribution area of the species. The two lowest strains are from southern Brazil, and the intermediate strains from Central America and from equatorial Brazil. Strains from localities geographically close to each other are more or less similar (Chilpancingo and Zopilote, Iporanga and Bertioga).

The discovery of negative isolation indices, which signify that matings between representatives of different strains are more frequent than between those of the same strain, was unexpected. Since sexual activity starts sooner after hatching of the flies from the pupae in Mexican than in Brazilian strains, it seemed possible that the results obtained in our experiments, in which flies of the same chronological age are placed together, reflected merely the difference in biological age of the flies. Indeed, females

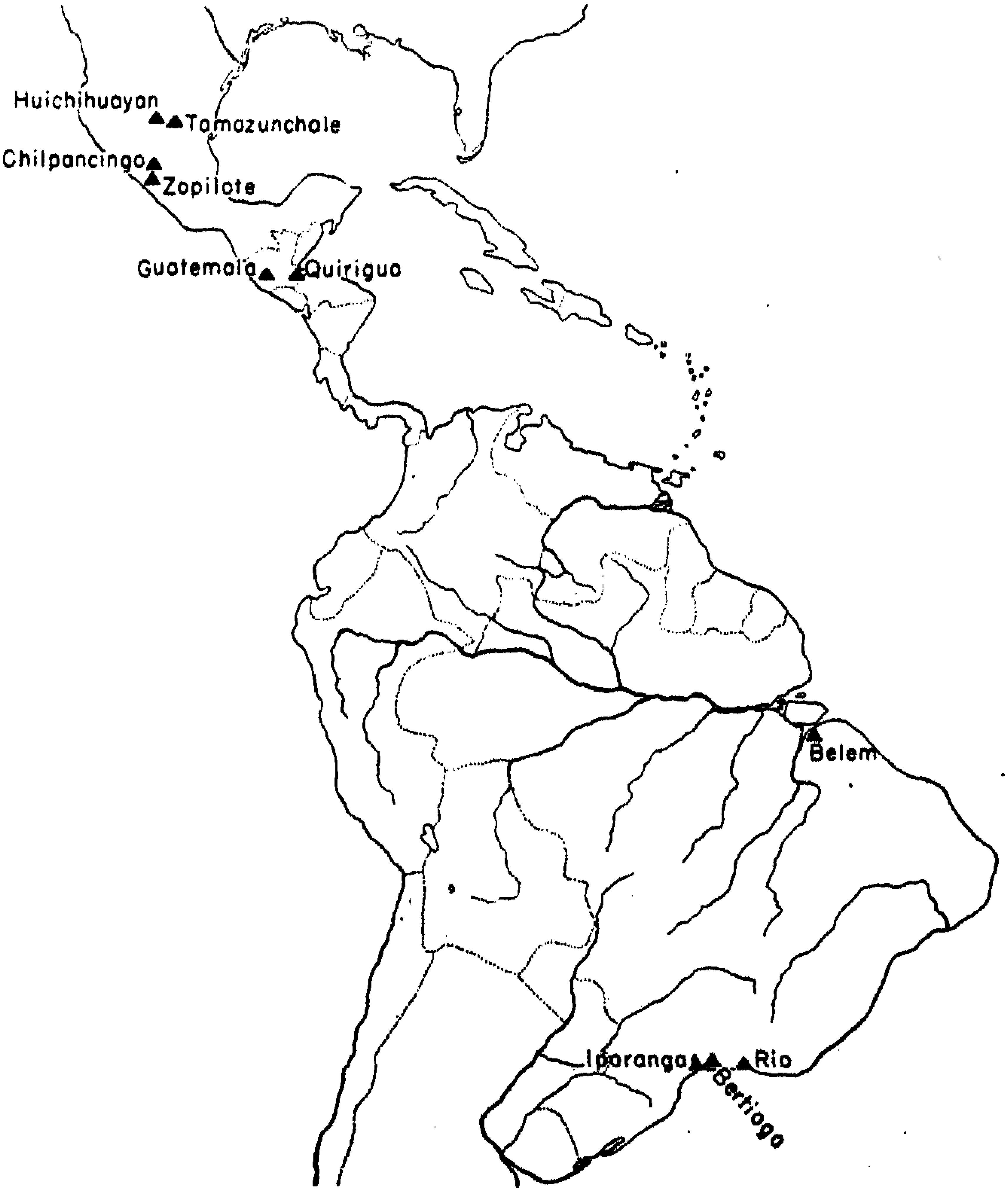


FIGURE 1

A map showing the localities in which strains of *D. prosaltans* and *D. sturtevanti* have been collected.

may become more sexually receptive as they grow older, and therefore older females may be inseminated in preference to younger ones. To test this possibility, Brazilian females were kept for two days on abundant food but without males, and then placed in vials with freshly hatched Mexican females and freshly hatched males of either kind. The females were dissected and examined as usual. The results are summarized in table 2. The rate of insemination clearly does not depend upon the age of the females alone, although the isolation indices in table 2 are lower than those for the corresponding crosses in table 1.

TABLE 2

Numbers of Females Dissected (*n*) and Per Cent Carrying Sperm (%) in Crosses in Which Brazilian Females Were Older Than the Mexican Ones

FEMALES	MALES	HOMOGAMIC *n*	HOMOGAMIC %	HETEROGAMIC *n*	HETEROGAMIC %	χ^2	ISOLATION INDEX
Chilpancingo, Bertioga	Chilpancingo	133	80.5	136	25.7	80.8	0.52
Chilpancingo, Bertioga	Bertioga	69	34.8	71	66.2	13.8	−0.31
Chilpancingo, Belem	Belem	41	53.7	40	80.0	5.3	−0.20

Summary.—The strains of *D. prosaltans* form a graded series with respect to their sexual preferences, and this seriation corresponds in a general way to the geographic origin of these strains. When northern (Mexican) males are kept with a mixture of northern and southern (Brazilian) females, more of the former than of the latter are inseminated. In similar experiments, the southern males also inseminate more northern than southern females. Females of the strains from Guatemala and from equatorial Brazil are discriminated against by South-Brazilian males in favor of females from the South-Brazilian strains.

---

[1] Dobzhansky, Th., and Mayr, E., these Proceedings, **30,** 238–244 (1944).

[2] Dobzhansky, Th., and Pavan, C., *Ibid.*, **29,** 368–375 (1943).

[3] This external difference is not sufficient to distinguish the Mexican and the South-Brazilian flies in all cases. Young Mexican flies have the same "faded-out" appearance as the older Brazilian flies, the "bright" coloration characteristic of the Mexican race developing only gradually with age. It does not seem advisable at this time to recognize the Mexican race taxonomically by attaching to it a Latin name.

[4] This condition, and the fact that the opening of the spermathecal duct is provided with a brush of chitinous needles protruding into the cavity of the spermatheca, suggests that the sperm stored in it is not used in fertilization of eggs. It seems possible that the function, in Drosophila, of the sperm in the chitinous spermathecae is to stimulate the development of the eggs in the ovary through a hormone-like mechanism.

[5] Stalker, H. D., *Genetics*, **27,** 238–257 (1942).

[6] It should be stressed that expressions like "preference" and "discrimination" are used merely to avoid circumlocutions. Our data show merely the frequencies of insemination of females of different strains, and it is not known whether these frequencies depend upon the behavior of males, of females or of both sexes.

MENDELIAN INHERITANCE OF ADAPTIVE ENZYMES IN YEAST*

By Carl C. Lindegren, S. Spiegelman and Gertrude Lindegren

The Henry Shaw School of Botany and the Department of Zoölogy, Washington University

Communicated September 13, 1944

Introduction. Our previous analysis[1, 2] of the mechanism of galactozymase production in *Saccharomyces cerevisiae* revealed that it is an adaptive[3] enzyme. The existence of unadaptable strains indicated that the initiation of enzyme synthesis was genetically controlled. The adaptation time for each specific strain was reproducible under standard conditions. While the genetic composition determines the initiation of enzyme synthesis, the quantitative level of enzyme in the cell below the maximum depends on an interaction between the specific fermentable substrate and the cytoplasm. It was thus of some interest to examine the adaptability of the progeny of two species differing in the possession of an adaptive enzyme.

S. validus and S. mandshuricus both possess the ability to ferment melibiose. Hybrids between these forms and other fermenting and non-fermenting yeasts led Winge and Laustsen[4] to state that "the presence of a specific enzyme is dominant to its absence in all the instances studied." They made matings by placing two spores in contact with each other. This method has the disadvantage that one cannot characterize the haplophase parents since both of the original spores are consumed in the mating. Furthermore, the genetical mechanism was not analyzed, since the phenotypes of the segregants from the hybrids were not examined. Finally, as will be shown here, melibiozymase is an adaptive enzyme, hence Winge and Laustsen's observations concerning the dominance of fermentative capacity need not imply a genetic basis. It is conceivable that hybrids of fermenting by non-fermenting strains may all be fermenters simply because they all carry 50% fermenting type cytoplasm (with its enzymes).

In the present experiments, the progeny of a hybrid between *S. carlsbergensis* and *S. cerevisiae* and other related progenies were examined. *S. carlsbergensis* can adapt to melibiose fermentation while *S. cerevisiae* cannot. The hybrids were produced by mixing haplophase cultures according to a method recently described.[5, 6, 7] Since only part of the culture is needed for the mating, the remainder can be used to determine the characteristics of the parent strain. Portions of the clone-culture are also available for back-crossing or mating to other clones of interest.

Methods.—One liter of the basic medium contained the following substances: 2 g. autolyzed yeast extract powder, 3 g. Bactopeptone, 1 g.

$(NH_4)_2SO_4$, 0.133 g. K_2PO_4, 1.866 g. KH_2PO_4, 0.25 g. $MgSO_4$, 0.125 g. $CaCl_2$, 1 cc. of 50% sodium lactate. Eight per cent cerelose was added to grow crops of yeast for adaptability test. Solid medium was made by adding 30 g. of agar to the above. A 12% solution of Eastman melibiose containing less than 0.01% of other fermentable carbohydrate was sterilized by Seitz filtration and added to the basic (carbohydrate free) medium to produce a solution containing 4% melibiose.

Test for Ability to Ferment Melibiose.—All 175 cultures were characterized according to their ability to ferment melibiose by the standard inverted tube technique. These tests were run in duplicate and over 90% of the cultures were held for a period of more than 30 days when the results were recorded. Many repetitions were made, also in duplicate, with uniformly identical results. The presence or absence of gas production is recorded in figure 2 by the signs + and –.

In using the inverted tube method, part of a clone is seeded into the broth containing melibiose as the carbohydrate source and allowed to grow. Since every mutation in a haplophase population becomes functional immediately, a positive test might not mean that the original clone possessed the fermentative capacity. Selection of a mutant produced during growth in the melibiose solution may have occurred. To exclude this possibility, these data were supplemented by duplicate tests of 12 critical cultures using Warburg manometers. A portion of the original clone was seeded into ordinary cerelose broth. After 48 hours' incubation, the resultant growth was centrifuged and washed twice with $M/15$ KH_2PO_4. It was then resuspended in $M/15$ KH_2PO_4. Melibiose was added in sufficient amounts to make a 4% solution. Under these conditions, selection is ruled out since no division takes place during the exposure to melibiose. If a stationary population exposed to melibiose acquires the ability to ferment the sugar, it can only be due to an interaction between the existing cells and the sugar. Samples were taken at intervals and tested for the ability to evolve CO_2 under anaerobic conditions. Fresh melibiose was always added to the suspension at the beginning of each determination to insure the presence of excess substrate.

It is highly improbable that a growing haplophase population kept in contact with cerelose (glucose) will acquire the ability to ferment melibiose because conditions would act against the selection of a melibiose-fermenting mutant. However, it is conceivable that a haplophase culture capable of fermenting melibiose grown for a long time on glucose might lose the ability to ferment melibiose by the selection of a non-fermenting mutant. There were no instances of this type during the present study, indicating that the melibiose fermenting genes are very stable. In all cases the results obtained with the manometric method agreed with those yielded by the inverted tube procedure.

Adaptation Times.—For accurate determinations of adaptation times, it was found[8] that the R. Q., which could be followed continuously by the two vessel method,[9] could be used. When the R. Q. reached a value of 1.8, the culture was considered adapted since this meant a $Q^N_{CO_2}$ value of 100. This was checked by flushing the KOH free cup with nitrogen after the R. Q. had exceeded 1.8 and measuring the $Q^N_{CO_2}$ value. All measurements were taken at 30.2°C., the temperature at which the suspensions were incubated with the substrate. The vessels were shaken at a rate of 100 oscillations per minute over a 7-cm. arc. The nitrogen used to displace the air in measurements of anaerobic CO_2 production was passed over hot copper to remove any traces of oxygen.

The Adaptive Nature of the Melibiose-Fermenting Enzyme.—To determine whether or not *S. carlsbergensis* is able to ferment melibiose immediately or only after a period of adaptation, phosphate suspensions of 48-hour, cerelose-grown cultures were placed in Warburg vessels with melibiose in the sidearms. After flushing with nitrogen, the melibiose was tipped. In no case was CO_2 evolved immediately or subsequently within 8 hours. However, if similarly prepared suspensions were incubated in air in the presence of melibiose and samples tested at intervals for anaerobic CO_2 evolution, measurable rates were found within 100 minutes, indicating that the enzyme-system necessary for melibiose fermentation is formed rapidly only after aerobic incubation in the presence of the substrate.

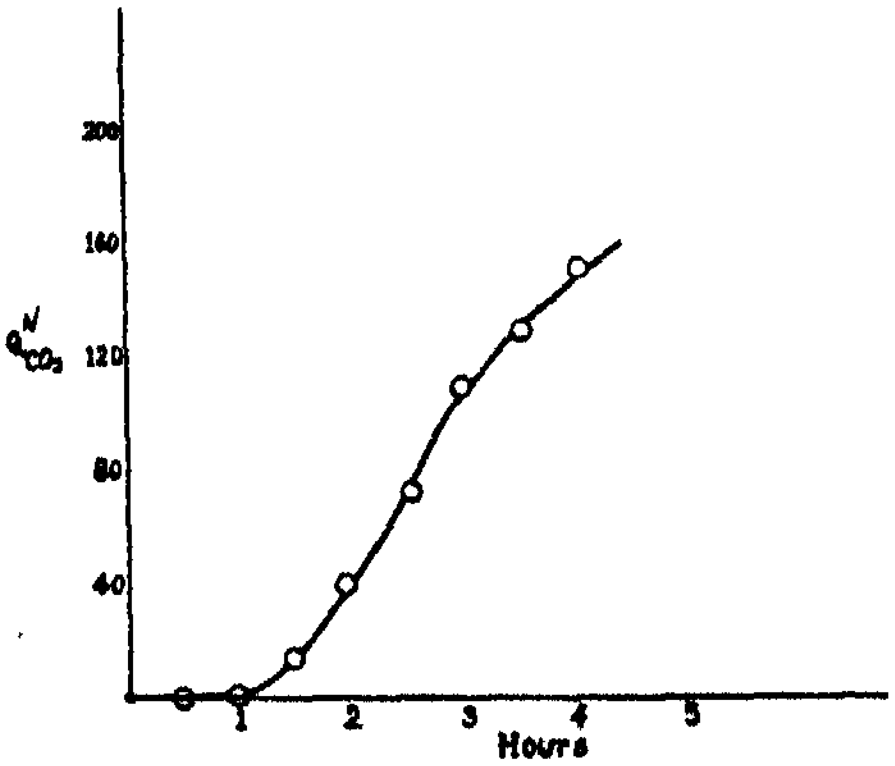


FIGURE 1

Graph showing adaptive nature of melibiose fermentation by *S. carlsbergensis.*

Figure 1 gives a representative curve of $Q^N_{CO_2}$ values after various periods of incubation. Four such tests were made in duplicate of the original cultures of *S. carlsbergensis* with the same results. Eight separate determinations of the adaptation times of this culture were made and an average value of 167 ± 10 minutes was obtained. It is clear from these experiments that the ability to ferment melibiose involves the formation of an adaptive enzyme. It differs in no essential way from the galactozymase activity previously studied. As in the case of galactose fermentation, cells which have been adapted to ferment melibiose lose this ability when removed from the substrate and have to be readapted to use it fermenta-

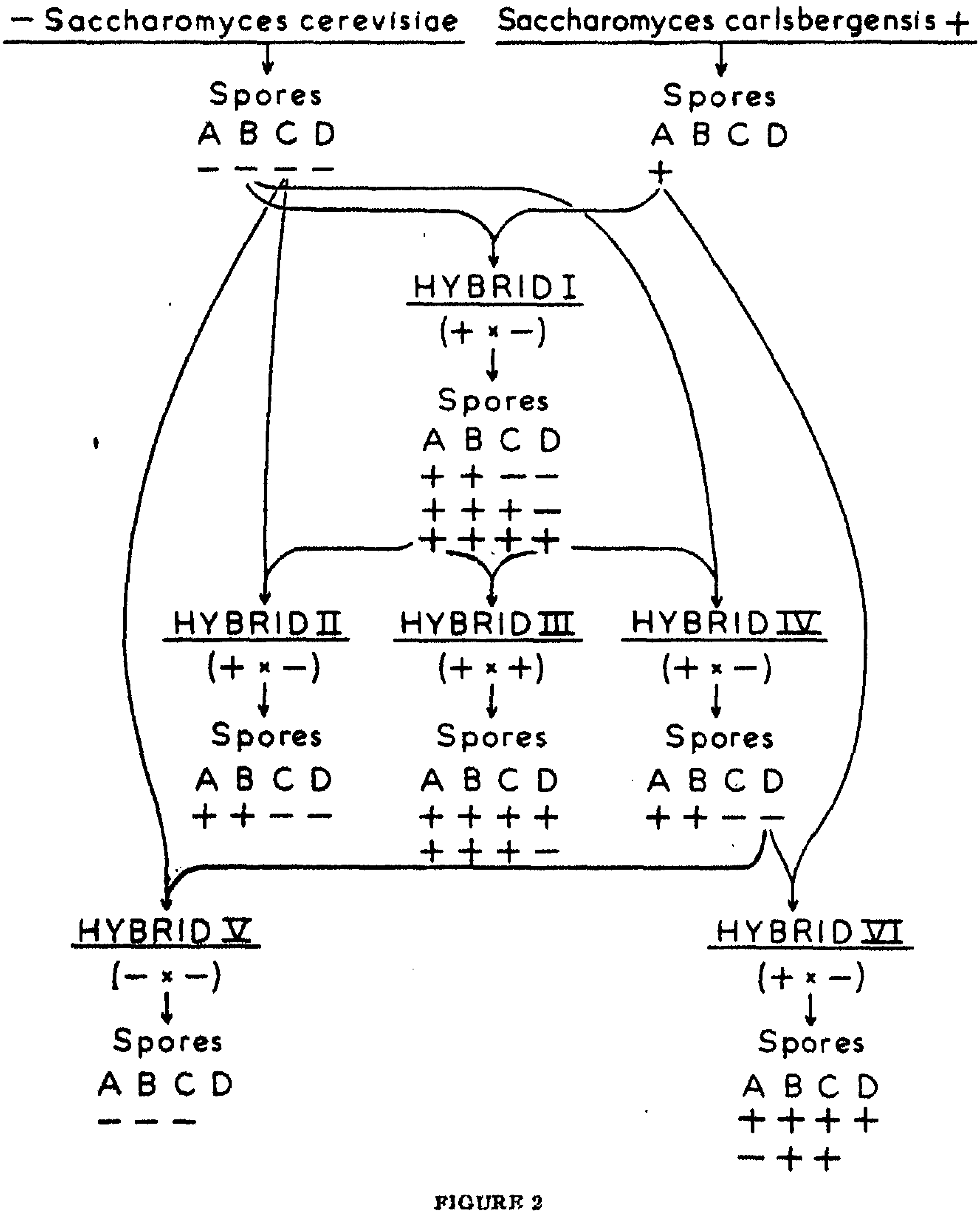


FIGURE 2

Fermentation of melibiose by diploid and and haploid progeny of *Saccharomyces cerevisiae* by *S. carlsbergensis* hybrids.

TABLE 1

FERMENTATION OF MELIBIOSE BY PROGENY OF VARIOUS CROSSES BETWEEN *Saccharomyces cerevisiae* (MELIBIOSE −) BY *Saccharomyces carlsbergensis* (MELIBIOSE +)

DIPLOID	PHENOTYPE OF DIPLOID	PHENOTYPES OF COPULANTS	NO. OF ASCUS	SPORE			
				A	B	C	D
S. carlsbergensis	+		C1	+			
			C5	+	+	+	+
			C6	+	+	+	+
S. cerevisiae	−		Lk5	−	−	−	−
Hybrid I	+	− x +	1	+	+	+	+
(Lk5B × C1A)			2	+	+	+	+
			3	+	+	+	+
			4	+	+	−	+
			5	+	+	+	
			6	+	−	+	+
			7	+	−	+	
			8	+	−	+	−
Hybrid II	+	− x +	1	+	+	−	−
(Lk5B × Hybrid I-1A)			2	+	−	+	
			3	+	−	+	−
			4	−		+	−
Hybrid III	+	+ × +	1	+	+	−	
(Hybrid I-1A × I-1D)			2	+	+	+	
			3	+	+	+	+
			4	+	−	+	+
			5	+	+	+	
			6		+	+	+
Hybrid IV	+	− × +	1	+		−	−
(Lk5C × Hybrid I-1D)			2	+	−		
			3	−	−	+	
			4	+	−	+	−
			5			−	+
			6		−		+
			7	−	+	+	−
			8	+	+	−	−
Hybrid V		− × −	1	−	−	−	
(Lk5C × Hybrid IV-7D)			2	−	−	−	
Hybrid VI	+	+ × −	1	+	+	+	+
(C1A × Hybrid IV-7D)			2	−	+	+	

Four spores were removed from each ascus and designated arbitrarily as *A, B, C, D*.

Pluses and minuses in these columns indicate that the spores, respectively, fermented or failed to ferment melibiose.

A blank space indicates that the spore failed to produce a viable colony.

tively. This adaptive behavior was true not only of the original *S. carlsbergensis*, but of all the haploid and diploid progeny which were tested, although the adaptation times varied. Wherever a clone is indicated as positive, it only showed that character after aerobic incubation in the presence of melibiose. None could ferment it on immediate contact.

Mendelian Segregation of the Adaptive Enzyme.—Figure 2 is a pedigree indicating the characteristic progeny of a hybrid between *S. cerevisiae* (mel −) and *S. carlsbergensis* (mel +), which are listed in detail in table 1. Four single ascospore cultures (Lk5A, B, C, D) were isolated from a single ascus of *S. cerevisiae*. Each one was haploid and all failed to adapt to melibiose fermentation. Several asci of *S. carlsbergensis* were also dissected, but in these early experiments only one spore per ascus survived. Each of these was adaptable. Two asci from *S. carlsbergensis* dissected later, gave eight cultures (C5A, B, C, D, and C6A, B, C, D) each of which proved adaptable.

Hybrid I was an interspecific hybrid (*cerevisiae by carlsbergensis* Lk5B × C1A) made by mixing melibiose + and melibiose − haplophase cultures. Three diploid cells isolated from this culture were all capable of fermenting melibiose, confirming the results of Winge and Laustsen. Eight asci were dissected from the interspecific hybrid and all the haplophase progeny were tested for the ability to ferment melibiose. The haplophase cultures from three asci all fermented melibiose. Two asci produced three + and one − culture. One ascus produced two + and two − cultures. The fact that in some cases more than two of the haplophase segregants from a single diploid nucleus fermented melibiose indicates that more than one gene is involved.

Hybrid II was produced by back-crossing positive haplophase culture I-1A to a negative haplophase culture from *S. cerevisiae*. A regular Mendelian segregation of the progeny shows that I-1A carries a single gene capable of controlling melibiose fermentation.

Hybrid IV was made by back-crossing positive haplophase culture I-1D to a negative haplophase culture from *S. cerevisiae*. In this case a regular Mendelian segregation again shows that a I-1D also carries a single gene.

Hybrid III was produced by mating I-1A by I-1D. Analysis of hybrids II and III has proved that each of these cultures carries a single gene controlling melibiose fermentation. If these genes are alleles, all the haplophase progeny of the hybrid should ferment melibiose. Two of the 20 haplophase segregants failed to ferment melibiose, proving that the gene carried by I-1A is not an allele of the gene carried by I-1D. Therefore, the original culture of *S. carlsbergensis* contained two different non-allelic loci controlling melibiose fermentation.

Hybrid V was made by back-crossing negative haplophase IV-7D to a negative haplophase from *S. cerevisiae*. Six haplophase progeny from two asci were all negative.

Hybrid VI was made by back-crossing the same culture (IV-7D) to a positive haplophase of *S. carlsbergensis*. Six of seven haplophase progeny fermented melibiose, while one failed. This finding together with the results obtained in Hybrid V confirms the fact that C1A possesses two genes controlling melibiose fermentation.

Summary and Conclusions.—*S. carlsbergensis* and its haploid segregants can adapt to melibiose fermentation while *S. cerevisiae* and its haplophase progeny cannot. Data obtained from 175 progenies of the interspecific and of related hybrids are consistent with the view that *S. carlsbergensis* contains two pairs of dominant genes, each of which controls the production of an adaptive enzyme. The gene initiates the synthesis of a melibiose fermenting enzyme complex, providing melibiose is present. *S. cerevisiae* is homozygous for the recessive alleles.

---

* Aided by a grant from Anheuser-Busch, Inc., St. Louis, Mo.

[1] Spiegelman, S., Lindegren, C. C., and L. Hedgecock, these PROCEEDINGS, **30**, 13–23 (1944).

[2] Spiegelman, S., and Lindegren, C. C., *Ann. Mo. Botan. Garden*, **31**, 219–233 (1944).

[3] Karstrom, J., *Ergeb. Enzymforsch.*, **7**, 350–376 (1937).

[4] Winge, O., and Laustsen, O., *Compt. rend. travaux lab. Carlsberg, série physiologique* **22**, 337–352 (1939).

[5] Lindegren, C. C., and Lindegren, G., these PROCEEDINGS, **29**, 306–308 (1943).

[6] Lindegren, C. C., and Lindegren, G., *Bot. Gaz.*, **105**, 304–316 (1944).

[7] Lindegren, C. C., and Lindegren, G., *Jour. Bact.*, **46**, 405–419 (1943).

[8] Spiegelman, S. in ms.

[9] Dixon, M., *Manometric Methods*, Cambridge, 1934, 122 pp.

APOMIXIS IN GUAYULE

BY KATHERINE ESAU

DIVISION OF BOTANY, COLLEGE OF AGRICULTURE, UNIVERSITY OF CALIFORNIA

Communicated September 25, 1944

The literature on the morphology of reproduction in guayule (*Parthenium argentatum* Gray) indicates that many irregularities occur in the development of the seed in this species. Thus Kirkwood[1] observed no division of the megaspore mother cell but in some preparations he found "what appeared to be vestiges of a non-functional spore." He suggested, therefore, that, at the most, the megaspore mother cell divided once before the 8-nucleate embryo sac was formed in the "usual manner." In her studies on the female gametophyte of guayule, Kokieva[2] listed a series of features that are usually taken as indications of parthenogenesis;

however, she was unable to prove that this phenomenon occurred in guayule. Dianowa, *et al.*,[3] also denied the existence of parthenogenesis in *P. argentatum* and other species of this genus.

The plants of *P. argentatum* that were used in the present study were of the 36- and 72-chromosome types. (Details regarding the origin of this material are to be published later.) The development of the female gametophyte and embryo differed in relation to the chromosome numbers. The 36-chromosome plants were largely, if not entirely, sexual; the 72-chromosome material was predominantly apomictic.

Table 1 illustrates the difference between the 36- and the 72-chromosome types with regard to the early development of the embryo sac. The figures in this table are based on an examination of young ovules ranging from 70 to about 400 microns in height. The youngest ovules showed megaspore mother cells enclosed within a nucellus that was still protruding beyond the integument; the oldest ovule was one with a tetranucleate embryo sac associated with a disintegrated nucellus and with an integument that completely enclosed the embryo sac and projected beyond it in the micropylar region.

The 36-chromosome material showed many megaspore mother cells in different stages of meiosis (item 3, table 1). Dyad and tetrad formation followed meiosis (items 5 and 6, table 1). Most of the uninucleate embryo sacs and many of the binucleate ones were associated with crushed megaspores. Sooner or later after the enlargement of the functioning megaspore the nucellus was destroyed. In the 72-chromosome material the high percentage of megaspore mother cells (items 2 and 3, table 1) and particularly of resting megaspore mother cells (item 2, table 1) indicated that the gametophyte formation proceeded slowly. The megaspore mother cells were of different sizes; some were so much enlarged that they were comparable to the uninucleate embryo sacs and were associated with a crushed nucellus (item 4, table 1). Views showing tetrads or crushed megaspores were few.

The occurrence of the relatively large, uninucleate structures associated with no crushed megaspores, but frequently with a disorganizing nucellus indicates omission of meiosis; conversely, the presence of megaspores (whether degenerating or not) is an evidence of preceding meiosis.[4] Judging by the items 4 to 6 in table 1 omission of meiosis is rare in the 36-chromosome plants used in this study but is common in the 72-chromosome plants. This evidence is strengthened by the observation that most of the megaspore mother cells in the 36-chromosome material were in some stage of meiosis, whereas most of those in the 72-chromosome plants were resting (items 2 and 3, table 1).

Thus apomixis in guayule involves *apomeiosis* of the type of *generative apospory*,[4] in which the megaspore mother cell does not form tetrads but directly develops into a uninucleate embryo sac with a diploid number of chromosomes. Since the apomictic 72-chromosome material showed some meiosis also the apospory in guayule may be termed *facultative.* Kirkwood[1] apparently noted this phenomenon in guayule but did not realize the significance of his observation.

The percentage of apospory in the 72-chromosome material can be roughly estimated from items 1 to 7 of table 1. Items 2 and 4 may be taken as indications of apospory and items 3 and 5 to 7 as evidences of meiosis. (Some of the megaspore mother cells in item 2 would probably have undergone meiosis; on the other hand, not all divisions in item 3 were identified as meiotic.) With these assumptions there was 74.2% of apospory in the 72-chromosome material. This figure agrees well with the percentages of maternal types (61.6 to 95.2%) that were observed by Powers and Rollins[6] in the F^1 generations of several *P. argentatum* Gray ♀ × *P. incanum* H.B.K. ♂ crosses. The same authors found no such apomictic reproduction when 36-chromosome plants were used as female parents and the present data suggest less than 4% apospory in similar plants.

TABLE 1

COMPARISON OF THE 36- AND 72-CHROMOSOME TYPES OF GUAYULE PLANTS WITH REGARD TO EMBRYO-SAC FORMATION

ITEM NUMBER	DESCRIPTION OF THE ITEM	36-CHROMOSOME TYPE	72-CHROMOSOME TYPE
1	Number of ovules examined	75	400
2	Per cent of resting megaspore mother cells	1.3	55.5
3	Per cent of dividing megaspore mother cells	21.4	12.0
4	Per cent of "uninucleate embryo sacs"[a] with disintegrating nucellus and no crushed megaspores	1.3	7.8
5	Per cent of dyads	5.4	1.7
6	Per cent of tetrads of megaspores	34.6	7.0
7	Per cent of uninucleate embryo sacs with crushed megaspores	8.0	1.3
8	Per cent of binucleate embryo sacs	8.0	12.2
9	Per cent of tetranucleate embryo sacs	20.0	1.8
10	Per cent of degenerated embryo sacs	0	0.7

[a] These are enlarged megaspore mother cells.

Generative apospory leads to the formation of an "unreduced gametophyte" with a diploid egg cell capable of giving rise to an embryo without fertilization. The occurrence of such embryo formation in guayule was checked by comparing pollinated and non-pollinated material from 36- and 72-chromosome plants. (The emasculations and pollinations were carried out by Dr. Reed C. Rollins of the Special Guayule Research Project, U.S.D.A.) The achenes from the non-pollinated flowers of a 36-chromosome plant contained no embryos; open pollinated plants of the same type contained embryos, usually associated with developing endosperm. Non-pollinated 72-chromosome plants contained embryos in most ovules, but the lack of pollination reduced the percentage of ovules in which the endosperm was developing and either reduced or completely inhibited the development of full-size embryos. In the pollinated and non-pollinated 72-chromosome plants there was a distinct lack of relationship between the early development of the embryo and that of the endosperm. The morphologic evidence on the beneficial effect of pollination upon embryo and endosperm development in the 72-chromosome material agrees with the data of Powers and Rollins[6] showing that pollination is necessary for the production of viable seed in the apomicts of guayule. The obvious conclusion is that in conjunction with the generative apospory *unreduced pseudogamy*[5] occurs in the apomictic types of guayule.

---

[1] Kirkwood, J. E., *Amer. Rev. Trop. Agri.*, **1**, 193–205 (1910).

[2] Kokieva, E., *Jour. Bot. U.R.S.S.*, **17**, 72–99 (1932).

[3] Dianowa, W. J., Sosnowetz, A. A., and Steschina, N. A., *Beih. Bot. Centralb.*, **53**, 294–308 (1935).

[4] Kiellander, C. L., *Svensk Bot. Tidskr.*, **31**, 425–429 (1937).

[5] Stebbins, G. L., Jr., *Bot. Rev.*, **7**, 507–542 (1941).

[6] Powers, LeRoy, and Rollins, Reed C., Manuscript submitted to *Jour. Amer. Soc. Agron.*

A METHOD IN RATIONAL DIOPHANTINE ANALYSIS

BY E. T. BELL

CALIFORNIA INSTITUTE OF TECHNOLOGY

Communicated September 25, 1944

1. The method furnishes rational solutions of a rational diophantine system when such solutions exist. By rational substitutions the system is reduced to a multiplicative system, which is completely solvable either in integers or in rational numbers. If the substitutions are birational, the complete integer or rational solution of the original system is obtained. When the substitutions are not rationally reversible, the method furnishes a partial solution of the original system; and this solution is rational in a certain necessary and sufficient number N of independent free parameters, the number being determined by the system. The obvious geometrical analog is the parametrization of a rational variety in space of a given number of dimensions; and the necessary and sufficient number N' of independent parameters is readily determined in any particular instance. That the arithmetical problem may be radically different from its geometric analog is suggested by the fact that, except in almost trivial systems, $N \neq N'$. For example, N' may be 9 while the corresponding N may exceed 40 million.

Although partial solutions are scarcely in the modern spirit of diophantine analysis, they sometimes imply by-products of interest in other parts of the theory of numbers. To cite an older instance, S. Ryley's[1] striking theorem (1825) that every rational number is a sum of three rational cubes, appears as a consequence of a partial solution of a very special ternary cubic diophantine equation. Geometrical considerations[2] have accounted for the success of Ryley's ingenious devices, which amount algebraically to the rational substitutions mentioned above. Actually, the special ternary cubic equation in question is replaced by a multiplicative system, whose complete solution is obtainable by the present method. But this solution is not the complete solution of the original equation, as the multiplicative system is not equivalent to the equation, an auxiliary restriction being introduced in order to reduce the equation to the multiplicative type. Similarly in all applications of the method to mixed additive and multiplicative systems; and the chief interest of the necessarily incomplete solutions is probably in their possible by-products. The method proposed is general and strictly elementary, and may be regarded as an alternative for the more recent applications of algebraic geometry to the arithmetic of rational varieties. It will suffice here to illustrate it by two examples, of which the first is the ternary cubic equation underlying Ryley's theorem.

2. To solve

$$(\xi + \eta + \zeta)^3 - d\xi\eta\zeta = m$$

in rational numbers ξ, η, ζ, where d, m are given rational numbers, the equation is reduced to the multiplicative type by assuming that $d\xi\eta\zeta + m$ is a product of linear homogeneous functions of the indeterminates ξ, η, ζ. This is equivalent to adjoining an auxiliary equation. A sufficient adjunction (one of an infinity) is essentially Ryley's,

$$d\eta\zeta^2 = m,$$

which reduces the equation to the multiplicative type

$$d\eta\zeta(\xi + \zeta) = (\xi + \eta + \zeta)^3,$$

which is of the form

$$dxyz = u^3.$$

For simplicity let d, m be integers. (If d, m are merely rational, the multiplicative equation is of the type $axyz = bu^3$ with a, b integers such that $(a, b) = 1$. The resolution is similar to that of the equation discussed, but the intermediate formulae involve more parameters.) The complete solution of the multiplicative equation is found in the usual way:

$$x = ay_1z_1x_1^2x_2^2x_3^3,$$

$$y = bx_2z_2y_1^2y_2^2y_3^3,$$

$$z = cx_1y_2z_1^2z_2^2z_3^3,$$

$$u = px_1x_2x_3y_1y_2y_3z_1z_2z_3,$$

in which the letters with suffixes are free parameters, and

$$a \equiv d_4d_6d_8d_3^2t,$$

$$b \equiv d_5d_6d_9d_1^2t,$$

$$c \equiv d_4d_5d_7d_2^2t,$$

$$p \equiv d_1d_2d_3d_4d_5d_6d_7d_8d_9d_{10}t$$

in which t is a free parameter and $d_1, \ldots, d_{10}$ are any integers such that

$$d = d_1d_2d_3d_4d_5d_6d_7^2d_8^2d_9^2d_{10}^3.$$

For given d the number of sets $(d_1, \ldots, d_{10})$ is finite and all sets may be found non-tentatively by solving a set of linear equations. With the above values of x, y, z we have

$$\xi + \eta + \zeta = u, \ \eta = x, \ \zeta = y, \ \xi + \zeta = z;$$

whence $\xi = z - y$, $\eta = x$, $\zeta = y$, subject to the condition

$$x + z = u.$$

The auxiliary equation is now

$$dxy^2 = m.$$

Values of the free parameters in x, y, z must be chosen so that the last two equations have rational solutions for two of the parameters in terms of the others. In this particular instance, z_1, z_2 are given immediately in the required form. (When such is not the case, if any of the parameters in the solution of the multiplicative equation occur with coprime exponents, an obvious replacement of these parameters by suitable powers of themselves, reduces the equation of condition and the auxiliary equation to the present special case.) On substituting the values of z_1, z_2 thus obtained into x, y, z, u, and then the resulting values into the expressions $\xi = z - y$, etc., for ξ, η, ζ we get a rational parametric solution of the given ternary cubic.

The solution may be greatly simplified by absorption: if certain of the parameters appear only as the same function of themselves wherever they occur, this function is no more general than a single free parameter, and may be replaced by one parameter. Here the absorption $x_1x_2^2x_3^2y_1^2y_2^3 \equiv \mu$

and the change in notation $z_3 \equiv \lambda$ followed by the substitution of $\mu\theta$ for λ, where μ, θ are free parameters, gives the solution in the form

$$\begin{aligned} D &\equiv a^2b^2d + cm\theta^3, \\ \xi &= (m^2p^6\theta^6 - D^3)/abdp\theta D^2, \\ \eta &= ma^2b^2dp^2\theta^2/D^2, \\ \zeta &= D/abdp\theta, \end{aligned}$$

where d, m are the given (integer) constants of the equation, and a, b, c, p are as stated above. For each set $(d_1, \ldots, d_{10})$ there is a solution ξ, η, ζ, and the number of such sets is finite. The identity $abcd = p^3$ is useful in verifying the solution.

The simplest set of values for the d_j is $d_1 = d$, $d_i = 1$, $i > 1$. For these $a = t$, $b = d^2t$, $c = t$, $p = dt$. Hence, with $t = \phi\theta$, there is the special solution

$$\begin{aligned} F &\equiv d^5\phi^3 + m; \\ \xi &= (m^2d^6\phi^3 - F^3)/d^4\phi^2F^2, \\ \eta &= md^7\phi^4/F^2, \\ \zeta &= F/d^4\phi^2, \end{aligned}$$

in which ϕ is a free parameter.[3]

3. The auxiliary equation reducing a given equation to the multiplicative type may render the equation algebraically reducible. Thus

$$(x + y)(y + z)(z + x) - dxyz = m,$$

with the auxiliary equation $m = dyz^2$ reduces the given equation to

$$y^2 + (x + az)y + xz = 0, \qquad a \equiv 1 - d,$$

on discarding a factor which yields only trivial solutions. For rational solutions, it is necessary and sufficient that the y-discriminant of this equation be a square,

$$(x + az)^2 - 4xz = w^2.$$

The last is a multiplicative equation,

$$(x + az + w)(x + az - w) = 4xz,$$

of which a particular solution is given by

$$\begin{aligned} x + az + w &= 2\alpha\theta, & x &= \alpha\phi, \\ x + az - w &= 2\beta\phi, & z &= \beta\theta, \end{aligned}$$

with α, β, θ, ϕ free parameters. There are thus two solutions,

$$\begin{aligned} w &= (\alpha^2 - 2\alpha\beta + a\beta^2)\theta/(\alpha - \beta), \\ x &= \alpha(\alpha - a\beta)\theta/(\alpha - \beta), \\ z &= \beta\theta, \\ y &= -\beta\theta(\alpha - a\beta)/(\alpha - \beta), \text{ or } -\alpha\theta. \end{aligned}$$

Imposing the auxiliary restriction on either of these, we get the rational solution of the original equation,

$$\begin{aligned} x &= -m(m + ad\theta^3)d\theta^2(m + d\theta^3), \\ y &= m/d\theta^2, \\ z &= \theta, \end{aligned}$$

in which $a = 1 - d$ and θ is a free parameter.

---

[1] Dickson, L. E., *History of the Theory of Numbers*, Vol. 2, 1920, pp. 726–729.

[2] Richmond, H. W., *Proc. Lond. Math. Soc.*, **21** (2), 401–409 (1922); *Jour. Lond. Math. Soc.*, **17**, 196–199 (1942). Mordell, L. J., *Ibid.*, 194–196. Richmond's papers contain references to the history.

[3] This solution appears to be distinct from those in the literature. However, it is not always easy to prove that formally distinct parametric solutions are numerically distinct.

GROUPS INVOLVING A SMALL NUMBER OF SETS OF CONJUGATE OPERATORS

By G. A. Miller

Department of Mathematics, University of Illinois

Communicated October 9, 1944

The identity constitutes a set of conjugate operators of every group. If a group G involves only one other set of conjugate operators all of its other operators are of the same order and hence it is a prime power group. Such a group always involves at least one invariant operator besides the identity and hence it results that G is then of order 2. If G contains three and only three sets of conjugate operators and is abelian it is the group of order 3. If it is non-abelian it cannot be a prime power group because a group of order p^m, p being a prime number contains at least $p^2 + p - 1$ sets of conjugate operators whenever it is non-abelian, as we proceed to prove.

Suppose that G is a non-abelian group of order p^m, p being an arbitrary prime number. The central quotient group of G is known to be non-cyclic and hence it is at least of order p^2, and the central of G is known to be at least of order p. If the central of G is of order p and the central quotient group of G is of order p^2 then each of the non-invariant operators of G is commutative with exactly p^2 operators of G and the number of these non-invariant operators of G is $p^3 - p$. The number of the sets of conjugate non-invariant operators of G is therefore $p^2 - 1$ and the number of the sets of conjugates of the invariant operators of G is p, therefore the number of the sets of conjugate operators of G is then $p^2 + p - 1$.

It has now been proved that every non-abelian group of order p^3, p being a prime number, contains exactly $p^2 + p - 1$ sets of conjugate operators. We proceed to prove that every other non-abelian group of order p^m contains more than $p^2 + p - 1$ sets of conjugate operators, so that the fact that a group of order p^m has $p^2 + p - 1$ sets of conjugate operators is a characteristic property of the non-abelian groups of order p^3. This fact results almost directly from the fact that the number of the sets of conjugate operators of a group is always larger than the number of the sets of conjugate operators in its quotient group with respect to any one of its non-identity invariant subgroups.

From what precedes it results that the smallest number of sets of conjugate operators in a prime power non-abelian group is five and that the two non-abelian groups of order 8 are the only prime power groups which separately contain exactly five sets of conjugate operators. It therefore results that if a group contains exactly three sets of conjugate operators and is non-abelian its order is divisible by at least two distinct prime numbers and when its order is divisible by exactly two distinct prime numbers it must be solvable. It must therefore contain an invariant subgroup of prime index and this index could clearly not exceed 2. The order of this invariant subgroup must be 3. That is, *the group of order 3 and the non-cyclic group of order 6 are the only two groups which satisfy the condition that all their operators appear in three and only three sets of conjugates under the group.*

To simplify some of the consideration which follow it may be desirable to formulate here the theorem that if a group G contains an invariant subgroup of prime index p which involves c sets of conjugate operators then the number of the sets of conjugate operators in G cannot be less than $p + (c - 1)/p$. This theorem results directly from the fact that in the quotient group there are at least $p - 1$ sets of conjugates and the number of the sets of conjugates in G resulting from the operators in the given invariant subgroup cannot be less than $(c - 1)/p + 1$. In the case of the theorem at the close of the preceding paragraph $p = 2$ and $c = 3$.

The only abelian groups which separately contain exactly four sets of conjugate operators are the two groups of order 4. It follows from the theorem noted above that if a group G is of prime power order it cannot contain exactly four sets of conjugate operators. If its order is divisible by two and only two distinct prime numbers it involves an invariant subgroup of prime index since it is solvable. This index is either two or three. If it is two, the order of the given invariant subgroup is odd, for if it were even two of the four sets of conjugates operators in G would be composed of operators of order 2 and hence all the operators of odd order in G would be conjugate under G.

This invariant subgroup of index two could not be of order 6 since the symmetric group of this order is a complete group. It therefore results that the given invariant subgroup of index two under G is the group of order 5 and G is the dihedral group of order 10. When the given invariant subgroup is of odd prime index under G it must be of index three and G is the tetrahedral group. That is, *if a group involves exactly four sets of conjugate operators it is one of the two groups of order 4 when it is abelian and when it is non-abelian it is either the dihedral group of order 10 or the tetrahedral group.* The latter group is simply isomorphic with the alternating group of degree four.

If a group contains exactly five sets of conjugate operators and is abelian it is the group of order 5. It was noted above that when a non-abelian prime power group contains exactly five sets of conjugate operators it is one of the two non-abelian groups of order 8. It should be emphasized that when the order of a group is divisible by a prime number p the number of its sets of conjugate operators cannot be less than the number of different orders of the divisors of p and $p - 1$ for the different such prime numbers. This theorem results directly from the fact that the group of isomorphisms of the group of order p is cyclic. Hence it results that if a group contains exactly five sets of conjugate operators its order cannot be divisible by a larger prime number than seven.

If the order of such a group is divisible by seven and contains only one subgroup of order 7 it is the dihedral group of order 14 or the semi-metacyclic group of order 21. In both of these cases the order of G is divisible by two and only two distinct prime numbers. In general, if its order is divisible by two and only two distinct prime numbers it must be solvable and hence it must contain an invariant subgroup of prime index. This index can obviously not exceed three and when it is three the corresponding invariant subgroup cannot be non-abelian since its central quotient group would then be non-cyclic and of order 4. It could not be the non-twelve group of order 24 since this group contains seven sets of conjugate operators.

The preceding considerations were greatly simplified by the fact that we were generally dealing with solvable groups. When the number of the sets of conjugate operators is five we can no longer assume that G is always solvable since the simple group of smallest composite order contains exactly five sets of conjugate operators. This group is of order 60 and is known to be simply isomorphic with the alternating group of degree five. One of the earliest workers in group theory, E. Galois (1811–1832), noted already that this group is simple and many of its other properties have received much attention since then.

SOME THEOREMS IN FINITE FIELD THEORY WITH APPLICATIONS TO FERMAT'S LAST THEOREM

By H. S. Vandiver

Department of Pure Mathematics, University of Texas

Communicated September 25, 1944

The first result we shall consider is as follows:

THEOREM I. *If the relation*

$$a_{m-1}x^{m-1} + a_{m-2}x^{m-2} + \ldots + a_0 = 0, \tag{1}$$

with coefficients in a finite field, F, has exactly k solutions x in F, and such that $x^m = 1$, then the cyclic matrix

$$\left|\begin{matrix} a_0 & a_1 \ldots a_{m-1} \\ a_{m-1} & a_0 \ldots a_{m-2} \\ \ldots\ldots & \ldots\ldots \\ \ldots\ldots & \ldots\ldots \\ a_1 & a_2 \ldots a_0 \end{matrix}\right| \tag{2}$$

is of rank $m - k$.

For a finite field of order p, a prime, and for $m = p - 1$, this result is due to König.[1] As our result may be proved by following closely the steps in the argument employed by Kronecker[1] (replacing $(p - 1)$ by $(p^n - 1)$ and $x^{p-1} = 1$ by $x^m = 1$), for proving König's theorem, the details of the proof will not be given here.

Dickson,[2] by extending a theorem due to Hurwitz,[3] obtained the following:

Let $f(x)$ be a polynomial, and let F be a finite field of order p^n which includes the coefficients. Then the number N of distinct roots $\neq 0$ of $f(x) = 0$, which belong to the field of order p^n, satisfies the relation

$$N + 1 \equiv C_0 + C_s + C_{2s} + \ldots \pmod{p},$$

where $s = p^n - 1$ *and*

$$f(x)^{p^n-1} \equiv C_0 + C_1x + C_2x^2 + \ldots$$

For $n = 1$, this becomes the Hurwitz[3] theorem.

We shall find it convenient to set up some results which are obtained by modification of the ideas of Hurwitz and Dickson. In the result just mentioned, Dickson generalizes Hurwitz's original notions by asking the number of solutions in any finite field which includes the coefficients of the original equation. We may go in the other direction and ask what solution our equation has within certain subsets of elements in the finite field $F(p^n)$ which includes the coefficients. For example, consider the polynomial, $f(x)$, with coefficients included in a $F(p^n)$, and suppose that ρ is a primitive root of $x^a = 1$ in $F(p^n)$; $p^n - 1 \equiv 0 (\bmod a)$. Let our equation be

$$f(x) = 0, \tag{3}$$

and consider

$$\sum_{t=0}^{a-1} (1 - f(\rho^t)^{p^n-1}). \tag{4}$$

This obviously equals, in the finite field, the number N of distinct solutions of equation (3) which lie in the set

$$1, \rho, \rho^2, \ldots \rho^{a-1}.$$

To simplify equation (4), write

$$(f(x))^{p^n-1} = C_0 + C_1x + \ldots. \tag{5}$$

We note that

$$\left.\begin{array}{ll} \sum_{h=0}^{a-1} \rho^{sh} = 0; & s \not\equiv 0 (\bmod a), \\ \sum_{h=0}^{a-1} \rho^{sh} = a; & s \equiv 0 (\bmod a). \end{array}\right\} \tag{6}$$

so that equation (4), using equation (5), gives

$$N = a - a(C_0 + C_a + C_{2a} + \ldots), \tag{7}$$

in the field.

The above method may be further extended. Let

$$f(x_1, x_2, \ldots, x_n) = 0 \tag{8}$$

where $F(p)^n$ is any finite field including the coefficients. Further, let ρ_i be a primitive root in the $F(p^n)$ of $= x_i^{a_i} = 1$;

$$p^n - 1 \equiv 0 (\text{mod } a_i),$$
$$i = 1, 2, \ldots\ldots, n.$$

We now seek all the distinct sets of solutions of (8) for $x_1, x_2, \ldots, x_n$, which are in the field $F(p^n)$ and which are also included among the ρ_i^k's. If N_n is this number, then we have

$$N_n = \sum(1 - (f(x_1, x_2, \ldots, x_n))^{p^n-1}) \tag{9}$$

where the summation extends over all the expressions obtained by letting x_i range independently over each of the elements

$$1, \rho_i, \rho^2{}_i, \ldots, \rho_i^{a-1};$$
$$i = 1, 2, \ldots, n.$$

Set

$$f(x_1, x_2, \ldots, x_n) = \sum C(e_1, e_2, \ldots, e_n) x_1^{e_1} x_2^{e_2} \ldots x_n^{e_n}$$

where the coefficients C are in $F(p^n)$. When x_i ranges over the ρ_i's; $i = 1, 2, \ldots, n$, we obtain, using equation (6)

$$\sum_{\rho} f(x_1, x_2, \ldots, x_n) = a_1 a_2 \ldots a_n \sum C(e_1' a_1, \quad e_2' a_2, \ldots, e_n' a_n)$$

where the summation on the right extends over all multiples $e_1' a_1$ of a_1 which are included among the e_1's, all multiples $e_2' a_2$ of a_2 included among the e_2's, and so on, for all the values of e_i; $i = 1, 2, \ldots, n$. Hence equation (9) becomes

$$N_n = A - A \sum C(e_1' a_1, e_2' a_2, \ldots, e_n' a_n) \tag{10}$$

where

$$A = \prod_{i=1}^{n} a_n.$$

We then have

THEOREM II. *If*

$$f(x_1, x_2, \ldots, x_n) = 0$$

is a polynomial, p^n *is the order of any finite field* $F(p^n)$ *including the coefficients of* f, *and* N_n *is the number of sets of solutions* $(x_1, x_2, \ldots, x_n)$ *such that each* x_i, $i = 1, 2, \ldots n$, *satisfies in the* $F(p^n)$, *the equation* $x^{a_i} = 1$, *where* $p^n - 1 \equiv 0 (\text{mod } a_i)$, *then, in the field,* N_n *is given by equation (10).*

We now obtain various applications of Theorems I and II to the theory of finite fields and to Fermat's last theorem. In another paper[4] the writer obtained the criteria

$$\binom{c}{s} + \binom{c}{s+l} + t^l + \binom{c}{s+2l} t^{2l} + \ldots \equiv 0 (\mathrm{mod}\ p) \qquad (11)$$

for the solution of

$$x^l + y^l + z^l = 0 \qquad (12)$$

where l is an odd prime, $y \equiv 0 (\mathrm{mod}\ l)$; $(xz, l) = 1$, p is a prime, $(p, l) = 1$, $(xyz, p) = 1$, and where, if $\mathfrak{p}$ is a prime ideal divisor of p in the field $k(\zeta)$, $\zeta = e^{2i\pi/l}$, then

$$c = \frac{N(\mathfrak{p}) - 1}{l}$$

and $\mathfrak{p}$ belongs to an exponent which is prime to l; $s = 1, 2, \ldots, l - 1$; $y/x = t$. Here also, as usual,

$$\binom{a}{b} = 0$$

if $b > a$.

Applying Theorem I of the present paper to the relation (2) of our former paper[4] we find that the matrices (11) and (11*a*) of the latter are of rank $\leqq c - 2$. Also if $(xyz, l) = 1$ then (11) and (11*a*) are of rank $\leqq c - 6$ for p sufficiently large

The criteria of Kummer for the solution of equation (12) in case I are as follows: *If* $(-t)$ *is any of the quantities* x/y, y/x, z/x, x/z, y/z, z/y, *then*

$$\left.\begin{aligned} B_n f_{p-2n-1}(t) &\equiv 0 (\mathrm{mod}\ l); \\ f_{p-2}(t) &\equiv 0 (\mathrm{mod}\ l); \\ n &= 1, 2, \ldots, \frac{l-3}{2}; \end{aligned}\right\} \qquad (13)$$

where $B_1 = {}^1/_2$, $B_2 = {}^1/_{30}$, etc., are the Bernoulli numbers and

$$f_a(t) = t + 2^a t^2 + \ldots + (l - 1)^a t^{l-1}.$$

If $s = p - 2n - 1$, then we examine the relation

$$f_s(t) \equiv 0 (\mathrm{mod}\ p). \qquad (14)$$

The above values for t show that this relation holds also for $1/t$, $(1 - t)/t$, $t/(1 - t)$, $1 - t$, $1/(1 - t)$ after using $x + y + z \equiv 0 (\mathrm{mod}\ l)$. (Note that t is here used in a different sense than in equation (11).) These six quantities are incongruent modulo l unless $t \equiv -1, +2$ or $-{}^1/_2$ or $t^2 - t + 1 \equiv 0$. The last combination is known to be inconsistent with equation (13) and the first holds only if $2^{l-1} \equiv 1 (\mathrm{mod}\ l^4)$. If equation (14) holds for six incongruent values then the cyclic matrix

$$\begin{vmatrix} 1^s & 2^s \ldots (l-1)^s \\ (l-1)^s 1^s & \ldots (l-2)^s \\ \cdots\cdots & \cdots\cdots \\ \cdots\cdots & \cdots\cdots \\ 2^s & 3^s \ldots \quad 1^s \end{vmatrix}$$

is of rank $\leqq l - 9$, since relation (14) is satisfied by $t \equiv \pm 1$ and $t \equiv 0$, and none of these lead to six incongruent values of t, modulo l, in fact $t \equiv 1$ is impossible from equation (12), and also $t \not\equiv 0$ by hypothesis.

We now apply Theorem II to the equation

$$au^m + bv^m + 1 = 0 \tag{15}$$

where $abuv \neq 0$, $mc = p^n - 1$, p prime, and seek the number N of its distinct solutions (u^m, v^m) in a finite field $F(p^n)$ including a and b. Here two solutions (u^m, v^m) and (u_1^m, v_1^m) are regarded as the same if and only if $u^m = u_1^m$ and $v^m = v_1^m$. Then Theorem II gives

$$N = \sum_{\rho_1, \rho_2} (1 - (a\rho_1 + b\rho_2 + 1)^{p^n-1}) \tag{16}$$

where ρ_1 and ρ_2 each range independently over the distinct roots of $z^c = 1$ in $F(p^n)$. Now, if $i = p^n - 1 - h$, then

$$(a\rho_1 + b\rho_2 + 1)^{p^n-1} = \sum_{h=0}^{p-1} (a\rho_1 + 1)^h (\rho_2 b)^i (-1)^h \quad .$$

since

$$\binom{p^n - 1}{h} \equiv (-1)^h (\text{mod } p).$$

Also

$$(a\rho_1 + 1)^h = \sum_{s=0}^{h} \binom{h}{s} (a\rho_1)^s.$$

Applying the last two relations to equation (16) gives

$$N = c^2 - c^2 \sum_{k=0}^{m} \sum_{r=0}^{k} \binom{kc}{rc} a^{rc} b^{mc-kc}$$

in the field. This reduces to a bit simpler form by noting that the term on the right corresponding to $k = 0$ is c^2, so that

$$N = -c^2 \sum_{k=1}^{m} \sum_{r=0}^{k} \binom{kc}{rc} a^{rc} b^{mc-kc}. \tag{17}$$

If $n = 1$, that is, the quantities a and b correspond to residue classes modulo p, then we may write

$$N \equiv -c^2 \sum_{k=1}^{m} \sum_{r=0}^{k} \binom{kc}{rc} a^{rc} b^{mc-kc} (\text{mod } p).$$

Let now $a = b = 1$, and $n = 1$, then equation (17) gives the result that N in equation (18) is now the number of solutions (u, v) of

$$u^m + v^m + 1 \equiv 0 (\text{mod } p) \tag{19}$$

where $(uv, p) = 1$; $p = 1 + cm$, which is related to Fermat's last theorem.

Return now to the consideration of equation (15). If we write $[h, k]$ for the number of distinct sets (s, t) such that

$$r^{sm+h} = r^{tm+k} + 1$$

in $F(p^n)$, r being a primitive root of F, it is known that[5] and

$$\sum_{i=0}^{m-1} [i, 0] = c - 1$$

$$\sum_{i} [i, j] = c \quad .$$

$i = 0, 1, 2, \ldots, m - 1$; $j = 1, 2, \ldots, m - 1$. Hence $[i, j] \leqq c$. For $n = 1$, then $c < p$, and we have

Theorem III. *The number of sets of distinct integral solutions* (u^m, v^m) *of*

$$au^m + bv^m + 1 \equiv 0 (\text{mod } p)$$

where a and b are integers, p is prime; $(abuv, p) = 1$*;* m *odd;* $p - 1 = mc$*; is the least residue,* $\geqq 0$*, of*

$$-c^2 \sum_{k=1}^{m} \sum_{r=0}^{k} \binom{kc}{rc} a^{rc} b^{mc-}$$

modulo p.

---

[1] Raussnitz, *Math. and Naturw. Berichte aus Ungarn*, **1**, 266–275 (1882–1883); Rados, *Jour. für Math.*, **99**, 258–260 (1886); Kronecker, *Vorlesungen über Zahlentheorie*, **1**, 389–415 (1901), including additions by Hensel (Kronecker refers to Königs' theorem in several places there as "sehr elegante" and "wichtige"). For other generalizations, see numerous references in Dickson's *History of Theory of Numbers*, vol. 1, pp. 226–233.

[2] *Bull. Amer. Math. Soc.*, **14**, 313 (1907–1908).

[3] *Archiv. Math. Phys.*, **5** (3), 17–27 (1903).

[4] *Proc. Nat. Acad. Sci.*, **28**, 144 (1942).

[5] Mitchell, *Ann. Math.*, **17**, 165 (1916). The trinomial equation (15) has been extensively studied. See references cited here by Mitchell, as well as Mitchell, *Ann. Math.*, **18**, 120 (1917). Dickson, *Am. Jour. Math.*, **57**, 391, 463 (1935); *Trans. Amer. Math. Soc.*, **37**, 363 (1935).

ON TRINOMIAL CONGRUENCES AND FERMAT'S LAST THEOREM

BY H. S. VANDIVER

DEPARTMENT OF PURE MATHEMATICS, UNIVERSITY OF TEXAS

Communicated September 25, 1944

In another paper[1] we obtained two theorems (numbered I and II) concerning finite fields, obtaining therefrom explicit expressions for the number of sets of solutions of certain equations which were applied, in special cases, to Fermat's last theorem. The present article deals with other results along these lines which involve trinomial congruences.

Applying Theorem III of F to Theorem IV of another paper[2] of the author's we obtain:

THEOREM I. *If it is possible to find a prime p such that $p = 1 + cl$; l an odd prime such that*

$$-c^2 \sum_{k=1}^{l} \sum_{r=0}^{k} \binom{kc}{rc} \tag{1}$$

is divisible by p and $c \not\equiv 0 (\text{mod } l)$ then

$$x^l + y^l + z^l = 0 \tag{2}$$

has no solution if $(xyz, l) = 1$.

Consider

$$au^m + bv^m + 1 \equiv 0 (\text{mod } p) \tag{3}$$

where all the symbols represent integers $(abuv, p) = 1$; p prime. This gives

$$(au^m + 1)^c \equiv (-1)^c b^c v^{mc} (\text{mod } p),$$
$$\equiv (-1)^c b^c (\text{mod } p),$$

using Fermat's Theorem and $mc = p - 1$. Expansion of the left-hand member gives, again using the theorem just mentioned,

$$\binom{c}{c-1} a^{c-1} u^{m(c-1)} + \binom{c}{c-2} a^{c-2} u^{m(c-2)} + \\ \ldots + \binom{c}{1} au^m + s \equiv 0 (\text{mod } p)$$

where

$$s = a^c + 1 + (-1)^{c+1} b^c.$$

We now apply Theorem I of the first paper to eliminate u^m and we find

$$\begin{vmatrix} a\,C_1\,a^2C_2 \ldots a^{c-1}C_{c-1}\,sC_0 \\ a^2C_2\,a^3C_3 \ldots\ldots sC_0\,aC_1 \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ sC_0\,aC_1 \ldots a^{c-2}C_{c-2}\,a^{c-1}C_{c-1} \end{vmatrix} \equiv 0,$$

modulo p, where

$$C_i = \binom{c}{i};$$

but this criterion looks much more complicated than the condition in Theorem III of F.

We shall now present still another method for attacking equation (3), which is due to the late H. H. Mitchell.

If $\theta = e^{2i\pi/p-1}$ then in the field $k(\theta)$ an ideal prime factor of (p) is $(\theta - v, p) = \mathfrak{p}$, where v is a primitive root of p in the rational field. Then

$$u^m \equiv v^{rm} \equiv \omega^r (\mathrm{mod}\ \mathfrak{p})$$
$$v^m \equiv v^{sm} \equiv \omega^s (\mathrm{mod}\ \mathfrak{p})$$

where $\omega = e^{2i\pi/c}$, and for some r and s. Hence equation (3) becomes

$$a\omega^r + b\omega^s + 1 \equiv 0 (\mathrm{mod}\ \mathfrak{p}).$$

Now take the norm of the left-hand member of this in $k(\omega)$; since said norm is rational it follows that

$$N(a\omega^r + b\omega^s + 1) \equiv 0 (\mathrm{mod}\ p). \qquad (4)$$

The left-hand member is now independent of p and l. It is clear then that unless we have an r and s such that this quantity reduces to zero then equation (3) is impossible for sufficiently large values of p. For the case $a = b = 1$, it is known that there is an r and s such that

$$\omega^r + \omega^s + 1$$

vanishes, if and only if $c \equiv 0$ (mod 3). Also

$$\begin{aligned} |\omega^r + \omega^s + 1| &\leqq |\omega^r| + |\omega^s| + 1 \\ &\leqq 1 + 1 + 1 \\ &\leqq 3 \end{aligned}$$

or

$$|N(\omega^r + \omega^s + 1)| \leqq 3^{\varphi(c)},$$

$\varphi(c)$ being the indicator of c, whence the result, due to Mitchell,[3] THEOREM II. *If c is a given integer* $\not\equiv 0 (\text{mod } 3)$, *and there is a prime p such that* $p = 1 + mc$ *and*

$$p > 3^{\varphi(c)}$$

then

$$u^m + v^m + 1 \equiv 0(\text{mod } p)$$

has no integral solutions (u, v); $(uv, p) = 1$.

It is known that (Dickson[4]) equation (3) is always satisfied in integers if $p = 1 + mc$, m is prime and

$$p \geqq (m - 1)^2(m - 2)^2 + 6m - 2$$

The question as to the congruence (5) having solutions or not for a prime p of the form $1 + mc$, which does not satisfy the last inequality or that given in Theorem II, is, in general, one of the great mysteries of number theory.

Employing Theorem II of this paper together with Theorem IV of another paper[2] we find

THEOREM III. *If c is a given even integer* $\not\equiv 0$ (*mod 3*) *and it is possible to find two primes l and p such that* $p = 1 + cl$; $l > c$; *and* $p > 3^{\varphi(c)}$; *then*

$$x^l + y^l + z^l = 0$$

is impossible in integers, with $(xyz, l) = 1$. *Here* $\varphi(c)$ *is the indicator of c.*

---

[1] These PROCEEDINGS, **30**, 362–367 (1944). This paper will be referred to as *F*.

[2] *Ann. Math.*, **27**, 56 (1926).

[3] This theorem and proof were communicated to me verbally by Prof. Mitchell. This happened some time in the year 1926, if I recall correctly. So far as I know he never published this material.

[4] *Jour. für die Reine u. Angew. Math.*, **135**, 134–141, 181–88 (1909).

ON THE INTEGRAL EQUATIONS OF CONTINUOUS DYNAMICAL SYSTEMS

By W. H. INGRAM

NEW YORK, N. Y.

Communicated July 5, 1944

1. *Introduction.*—The dynamical systems are the one-dimensional ones considered recently[1] and are governed by the vector equation

$$\Upsilon(x, t): = \int_a^b \mathcal{G}(x, \xi)\mathcal{A}(\xi) \frac{\partial \Upsilon(\xi, t):}{\partial t} d\xi \qquad (1)$$

in which the components of Υ: are the coördinates, velocities and internal forces of the system and where $\mathcal{G}$ is the matrix of the Green's functions and $\mathcal{A}$ the matrix of the conservative parameters (inertial and elastic coefficients).

If the variables are separable, as it is necessary to suppose, there exist relations of the form $\Upsilon_i(x, t) = W_i(x)T_i(t)$, $i = 1, 2, \ldots, n$ for the components of the vector in equation (1) and if we define B_{ij} by the equation

$$\int_a^b K_{ij}(x, \xi)W_j(\xi)d\xi = B_{ij}W_i(x)$$

in which K_{ij} is the element of $\mathcal{G}\mathcal{A} = \mathcal{K}(x, \xi)$ in the ith row and jth column, then the time factors are components of the vector T: which satisfies the equation $\mathcal{B}\dot{T}: = T$: in which $\mathcal{B} = (B_{ij})$. This equation in T: has the solution $T: = \epsilon^{\mu t}C$: provided C: and μ satisfy the equation $\mu\mathcal{B}C: = C$:. When a solution for C: of this equation exists, the substitution $C_iW_i(x) = \epsilon^{-\iota\beta_i}U_i(x)$ into equation (1) gives the integral equation

$$\mathcal{E}U(x): = \mu\int_a^b\mathcal{G}(x, \xi)\mathcal{A}(\xi)\mathcal{E}U(\xi):d\xi$$

in which $\mathcal{E}$ is the constant diagonal matrix[2] $[\epsilon^{-\iota\beta_i})$, for the determination of the β's, $U(x)$: and μ. On the hypothesis that the μ's are distinct, a necessary and sufficient condition that the variables be separable is that the B's be constants.

This integral equation and its *associate*, written in the form

$$\Phi(x): = \mu\int_a^b\mathcal{K}(x, \xi)\Phi(\xi):d\xi, \qquad \cdot\Psi(x) = \mu\int_a^b\cdot\Psi(\xi)\mathcal{K}(\xi, x)d\xi \quad (2)$$

where $\mathcal{K}$ is in general a non-symmetric matric kernel, are to be found, in effect, in Fredholm's classic paper[3] but, as is well known, development of the theory has remained seriously incomplete, when the kernel is not restricted to be symmetric, and notably in directions admitting any aphoristic geometrical restatement of the results. In this regard the theory of equation (2) is in remarkable contrast to the theory of the *conjugated* pair of equations

$$\phi(x) = \kappa\int_a^b K(x, \xi)\psi(\xi)d\xi, \qquad \psi(x) = \kappa\int_a^b\phi(\xi)K(\xi, x)d\xi$$

of Erhard Schmidt[4] which has a complete analogy in the abstract theory of Hilbert space but, nevertheless, slight and insignificant contact with the differential equations of dissipative elastokinetics.

Directly connected with the differential equations of such dynamical systems and their adjoint equations, on the other hand, is the *adjoint* pair of integral equations

$$Y(x): = \mu\int_a^b\mathcal{G}(x, \xi)\mathcal{A}(\xi)Y(\xi):d\xi, \qquad \cdot Z(x) = \mu\int_a^b\cdot Z(\xi)\mathcal{A}(\xi)\mathcal{G}(\xi, x)d\xi$$

studied by Birkhoff & Langer[5] who obtained formal results of a kind indicated by the theory of the affine transformation of an n-dimensional space into itself and among them the bilinear expansion $\mathcal{G}(x, y) = \Re\Sigma\lambda_r Y_{(r)}(x) :\cdot Z_{(r)}(y)$, $\lambda_r = \mu_r^{-1}$ for the Green's matrix and equivalent to the expression $\mathcal{K}(x, y) = \Re\Sigma\lambda_r\Phi_{(r)}(x) :\cdot \Psi_{(r)}(y)$ for the kernel $\mathcal{K}$.

There is evidence that there exists an important class of kernels the members of which either have this expansion or their iterates have a similar one. In fact, the representation for the iterated kernel

$$\mathcal{K}^s(x, y) = \Re\Sigma \frac{\Phi_{(r)}(x) :\cdot \Psi_{(r)}(y)}{\mu_r^s} \tag{3}$$

is known to be valid in the cases of the vibrating string and of the electrical transmission line, in the usual cases where the Green's function satisfies the equation $d^2G(x, y)/dx^2 = 0$, and for the non-symmetric kernel in the integral equation for the thermo-mechanical motions of an elastic rod[6] with $s = 1$. It is known that the series converges absolutely and uniformly in both variables to the kernel or iterated kernel with $s = 1$ if the kernel is symmetric, continuous and positive definite (Mercer), with $s = 2$ if the kernel is symmetric and continuous (Erhard Schmidt). The series is also known (Garbe,[7] Mercer[8]) to converge to a non-symmetric kernel of the form $K(x, y) = \alpha(x)k(x, y)$ for $k(x, y)$ "allgemein" or for $K(x, y)$ "completely symmetrisable" under certain conditions. The modal numbers of a real, symmetric and segmentally continuous kernel are either finite in number or enumerably infinite and converge to ∞[9] and the same is true for the non-symmetric and only segmentally continuous matric kernel of the electrical transmission line. In fact, the literature of integral equations and of elastokinetics would appear to contain very few if any examples where the modal numbers have or could have any finite cluster point.

In view of this evidence it is obvious that the class of kernels for which equation (3) is in some sense valid and for which the modal numbers have no finite cluster point is sufficiently extensive and important to warrant such steps forward as are possible on the basis of these properties even although the elusive problem of characterization of the class remains unsolved. The step essayed in the following paragraphs is in the direction of a more explicit determination of the modal vectors and numbers.

Notation.—The notation used in this paper is a modification of that of Birkhoff & Langer. A line or row vector $(v_1, v_2, \ldots, v_n)$ is written $\cdot v$ but a column vector of elements u_i, $i = 1, 2, \ldots, n$ is written $u:$. The inner product $\Sigma v_i u_i$ of a row vector $\cdot v$ and a column vector $u:$ is written $\cdot vu:$ and the outer product $(u_i v_j)$ is an $n \times n$ matrix with the element $u_i v_j$ in the ith row and jth column and is written $u :\cdot v$. Script letters denote $n \times n$ matrices. As powers of the kernel matrix $\mathcal{K}$ do not occur in the work, the s-fold iterated integral

$$\int_a^b \int_a^b \ldots \int_a^b \mathcal{K}(x, \xi_1)\mathcal{K}(\xi_1, \xi_2)\ldots\mathcal{K}(\xi_s, y)d\xi_1\, d\xi_2\ldots d\xi_s$$

is written simply $\mathcal{K}^{s+1}(x, y)$.

2. *Determination of the Modal Vectors and Numbers.*—The method for the computation of the first modal number and function of any real symmetric kernel first given by Erhard Schmidt and the extension of the method by Koch to the cases of the second, third, ... modal number and function in succession have been still further generalized by Ingram[1] to the case of a vector integral equation with non-symmetric matric kernel. Noteworthy steps in this generalization are the recognition of the existence of possible differences in phase among the various components of any modal vector $\Upsilon(x, t)$:, the simultaneous computation of the Φ's and Ψ's, the matric concepts involved and the algebraic problem which is the end-product of the method. Certain sufficient conditions for the convergence of the generalized procedure are now considered.

The relations of orthogonality

$$\left.\begin{array}{r} \int_a^b \cdot\Psi_{(i)}(x)\Phi_{(j)}(x):dx = 0 = \int_a^b \cdot\Psi_{(i)}(x)\Phi_{(j)}(x):dx \qquad i \neq j \\ \int_a^b \cdot\Psi_{(i)}(x)\Phi_{(i)}(x):dx = 0 \qquad \mu_i \neq \bar{\mu}_i \qquad . \\ \int_a^b \cdot\Psi_{(i)}(x)\Phi_{(i)}(x):dx = 2 \end{array}\right\} (4)$$

between the modal functions are first derived from equation (2) in the usual way: the last equation is a matter of agreement.

LEMMA: *If the series* $\Re\Sigma\lambda_r^s\Phi_{(r)}(x):\cdot\Psi_{(r)}(y)$ *is segmentwise uniformly and boundedly convergent (with respect to* y*) to* $\mathcal{K}^s(x, y)$ *for every fixed value of* x *on* $a \leqslant x \leqslant b$ *and for some natural number* s*, if* $\Xi(x)$: *is any integrable vector orthogonal to* $\cdot\Psi_{(1)}(x)$, $\cdot\Psi_{(2)}(x)$, ..., $\cdot\Psi_{(\kappa-1)}(x)$ *but not to* $\cdot\Psi_{(\kappa)}(x)$ *and such that* $\Sigma\int_a^b \cdot\Psi_{(r)}(y)\Xi(y):dy$ *converges absolutely, and if the modal numbers can be enumerated in order of increasing modulus and have no finite accumulation point, then the vectors*

$$\mathfrak{P}^s(x): = \int_a^b \mathcal{K}^s(x, y)\Xi(y):dy, \qquad \mathfrak{P}^{s+1}(x): = \int_a^b \mathcal{K}(x, y)\mathfrak{P}^s(y):dy$$

are either constant multiples of one and the same function of x *or become so as* $s \to \infty$.

Proof. The first step is to make it clear that both sides of equation (3) may be multiplied by $\Xi(y)$: and integrated term-by-term with respect to y so as to give

$$\mathfrak{P}^s(x): = \int_a^b \mathcal{K}^s(x, y)\Xi(y):dy = \Re\Sigma\lambda_r^s\tau_r\Phi_{(r)}(x):, \ \tau_r = \int_a^b .\Psi_r(y)\Xi(y):dy$$

an equation in which the series converges. This result follows from the definition of a uniformly and boundedly convergent series and from the fact that a segmentwise uniformly and boundedly convergent series of terms which are functions of y may be multiplied term-by-term by any integrable function and integrated term-by-term to give the integral of the product of the sum limit multiplied by the function.

It is to be noted next that the modal functions are bounded since the kernel is bounded from the definition of the Green's matrix and because $\mathfrak{A}(x)$ is naturally bounded (but possibly discontinuous at any number of points) and accordingly any modal function which is finite for one value of x is finite for all values on $a \leqslant x \leqslant b$. If $M\colon$ is a vector of positive elements such that $|\Phi_{(r)i}(x)| \leqslant M_i$, all i, then the last equation becomes the inequality

$$|\mathfrak{P}^s(x)|\colon \leqslant M\colon\{|\lambda_\kappa^s\tau_\kappa| + |\lambda_{\kappa+1}^s\tau_{\kappa+1}| + \ldots\}.$$

But by hypothesis $\tau_\kappa + \tau_{\kappa+1} + \ldots$ converges absolutely and the moduli of the λ's form a decreasing sequence and tend to zero as a limit for $s = 1$ and tend to zero faster for $s > 1$. For sufficiently large s, therefore, the first non-zero term can be taken as an approximation to the series with an error for any value of x as small as desired short of zero. It follows that $\mathfrak{P}^s(x)\colon$ and $\mathfrak{P}^{s+1}(x)\colon$ are both approximately proportional to the same real vector $\Re\lambda_\kappa^s\tau_\kappa\Phi_{(\kappa)}(x)\colon$ which is seen to be a diagonal matrix times the real vector $U_{(\kappa)}(x)\colon$.

The lemma just proved justifies a procedure for a more explicit determination of the modal functions and numbers. If, in conformity with the equation $\mathcal{B}\dot{T}\colon = T\colon$ satisfied by the time factor in the variables of equation (1), the substitution $\Upsilon_i(x, t) = \epsilon^{\alpha t}\cos(\omega t - \beta_i)U_i(x)$ is made, then equation (1) takes the form $\Phi(x)\colon = \mu\int_a^b\mathcal{K}(x, \xi)\Phi(\xi)\colon d\xi$ with, as before $\mu = \alpha + \iota\omega$, $\Phi(x)\colon = \mathcal{E}U(x)\colon$, $\mathcal{E} = [\epsilon^{-\iota\beta_i})$. By the lemma, $\int_a^b\mathcal{K}^s(x, \xi)\Xi(\xi)\colon d\xi$ gives, at least approximately, the components $\Phi_{(r)1}(x), \Phi_{(r)2}(x), \ldots, \Phi_{(r)n}(x)$ each to within a constant factor. It follows directly, if, as a matter of notation $\Phi(x)\colon = W(x)\colon - \iota X(x)\colon$, that there are the at least approximate equations

$$W_i(x) = \eta_i\mathfrak{P}_i^s(x), \qquad X_i(x) = \eta_i \tan\beta_i\mathfrak{P}_i^s(x) \tag{5}$$

where η_i is a constant factor and in which all quantities are real when, as is always possible, $\Xi(x)\colon$ is taken real. Substitution of these expressions into equation (3) gives

$$\left.\begin{aligned} \eta_i\mathfrak{P}_i^s(x) &= \alpha\Sigma\eta_j\int_a^b K_{ij}(x, \xi)\mathfrak{P}_j^s(\xi)d\xi \\ &\qquad + \omega\Sigma\eta_j\tan\beta_j\int_a^b K_{ij}(x, \xi)\mathfrak{P}_j^s(\xi)d\xi \\ \eta_i\tan\beta_i\mathfrak{P}_i^s(x) &= -\omega\Sigma\eta_j\int_a^b K_{ij}(x, \xi)\mathfrak{P}_j^s(\xi)d\xi \\ &\qquad + \alpha\Sigma\eta_j\tan\beta_j\int_a^b K_{ij}(x, \xi)\mathfrak{P}_j^s(\xi)d\xi \end{aligned}\right\} \tag{6}$$

a system of $2n$ equations in the unknowns $\alpha, \omega, \eta_1, \eta_2, \ldots, \eta_n, \tan\beta_1, \tan\beta_2, \ldots, \tan\beta_n$. When $\mathfrak{P}_i^s(x)$ and $\int_a^b K_{ij}(x, \xi)\mathfrak{P}_j^s(\xi)d\xi$ are multiples of the same function of x, all i and j, i.e., when $\int_a^b K_{ij}(x, \xi)\mathfrak{P}_j^s(\xi)d\xi = p_{ij}\mathfrak{P}_i^s(x)$ with p_{ij} a constant, then there is the system of $2n$ equations

$$\left.\begin{aligned} \eta_i &= \quad \alpha\Sigma p_{ij}\eta_j + \omega\Sigma p_{ij}\eta_j \tan\beta_j \\ \eta_i \tan\beta_i &= -\omega\Sigma p_{ij}\eta_j + \alpha\Sigma p_{ij}\eta_j \tan\beta_j \end{aligned}\right\} \quad (7)$$

to be solved for α, ω, the $n - 1$ ratios η_i/η_1 and $n - 1$ of the β's and in which the p's are known and one β is arbitrary. The ultimate validity of equation (7) is ensured when $\mu_1\bar{\mu}_1 < \mu_2\bar{\mu}_2 < \ldots$ and the μ's have no finite accumulation point and the p's are constants for then any error in equation (5) tends to zero as $s \rightarrow \infty$.

The constancy of the p's is connected with our fundamental assumption that the variables in equation (1) can be separated. Obviously, if the B's of the Introduction are constants then any μ must be a constant and the corresponding T: cannot be a function of x. The constancy of the B's is also a necessary condition for any T: to be a function of t only. For let $\mathfrak{J} = [T_1\colon T_2\colon \ldots T_n\colon]$ be a matrix of the fundamental solutions of the equation $\mathcal{B}\dot{T}\colon = T\colon$. By hypothesis, the solutions of the equation $\det(\mu\mathcal{B} - \mathcal{I}) = 0$ are distinct so that $\det \mathfrak{J} \neq 0$. Hence $\mathcal{B} = \mathfrak{J}(\dot{\mathfrak{J}})^{-1}$ and no element of $\mathcal{B}$ can be a function of x. As has been seen in the proof of the lemma, $\mathfrak{P}_i^s(x)$ is approximately proportional to $U_{(\kappa)i}(x)$ so that the ultimate constancy of the p's follows. When the variables cannot be separated, there are no modal functions and numbers.

The system (7) obviously is a linear one in the unknowns $\eta_1, \eta_2, \ldots, \eta_n, \eta_1 \tan\beta_1, \eta_2 \tan\beta_2, \ldots, \eta_n \tan\beta_n$ and so the problem may be expressed as that of solving the vector system

$$\left.\begin{aligned} \alpha\mathcal{P}\eta\colon + \omega\mathcal{P}\zeta\colon &= \eta\colon \\ -\omega\mathcal{P}\eta\colon + \alpha\mathcal{P}\zeta\colon &= \zeta\colon \end{aligned}\right\} \quad (8)$$

in which, as a matter of notation, $\zeta_i = \eta_i \tan\beta_i$ and $\mathcal{P} = (p_{ij})$. Real values of α and ω therefore exist which will reduce the rank of the matrix

$$\mathfrak{M} = \begin{bmatrix} \alpha\mathcal{P} - \mathcal{I} & \omega\mathcal{P} \\ -\omega\mathcal{P} & \alpha\mathcal{P} - \mathcal{I} \end{bmatrix}$$

to $2n - 2$. The explicit formulae for the unknowns which accomplish this reduction in the case $n = 2$ have been applied to the case of the electrical transmission line.[1] The discussion is summarized in the following theorem.

THEOREM 1. *The hypotheses are* (1) *the modal numbers for equation* (2) *occur only in conjugate complex pairs such that* $\mu_1\bar{\mu}_1 < \mu_2\bar{\mu}_2 < \ldots$ *and have no finite accumulation point,* (2) $\mathcal{K}^s(x, y)$ *can be expressed, for some s, as a finite sum or as a stepwise uniformly and boundedly convergent series of the form* (3), (3) $\Xi(x)\colon$ *is a linear combination of* $\Phi_{(\kappa)}(x)\colon, \Phi_{(\kappa+1)}(x)\colon, \ldots$ *or has an absolutely and uniformly convergent representation in terms of these functions. The conclusions are, with* $\Xi(x)\colon$ *taken real, that there exist real values* *of* α *and* ω *which will reduce the rank of* $\mathfrak{M}$ *to* $2n - 2$, *that such values are at least approximately the real and imaginary parts, respectively, of the modal number* μ_κ *for the equation* (2) *and that the corresponding solution of the equation* $(\cdot\eta, \cdot\xi)\mathfrak{M}_T = 0$ *with one element each of the vectors taken arbitrarily gives, in the equations*

$$W_{(r)}(x){:} = [\eta)\int_a^b \mathcal{K}^s(x, y)\Xi(y){:}\,dy, \qquad X_{(r)}(x){:} = [\zeta)\int_a^b \mathcal{K}^s(x, y)\Xi(y){:}\,dy$$

the modal function $\Phi_{(r)}{:}$.

3. *The Subsidiary Modal Determinant.*—Suppose there is a solution $\Phi(x){:}$ of equation (2). Then $\mathfrak{N}\Phi(x){:}$ with $\mathfrak{N}$ a diagonal matrix of complex numerical elements is also a solution provided only that the columnar contraction $N{:}$ of $\mathfrak{N}$ be a solution of the equation $\mu\mathcal{B}N; = N{:}$ in which the elements of $\mathcal{B}$ are defined by the equation $\int_a^b K_{ij}(x, \xi)\Phi_j(\xi)d\xi = B_{ij}\Phi_i(x)$ equivalent to the definition in the Introduction. In fact, every solution of the equation $\mu\mathcal{B}N{:} = N{:}$ for μ and corresponding $N{:}$ gives a solution of equation (2), the proof being by contradiction.

If, instead of $\Phi(x){:}$, we take $\mathfrak{N}_1\Phi(x){:}$ such that $\mathfrak{N}_1{:}$ is any one of the solutions of the subsidiary equation $\mu\mathcal{B}N{:} = N{:}$ and such that μ' is the corresponding modal number and suppose $\mathfrak{N}\mathfrak{N}_1\Phi(x){:}$ to be a modal vector having the modal number μ'' not necessarily the same as either μ' or μ, then reasoning parallel to that just given shows that the vector $(NN_1){:}$ which is the columnar contraction of $\mathfrak{N}\mathfrak{N}_1$ satisfies the equation $\mu\mathcal{B}(NN_1){:} = (NN_1){:}$ and so the set of $\mathfrak{N}$'s have closure under multiplication.

THEOREM 2. *If* $\Phi(x){:}$ *is any modal vector, then* $\mathfrak{N}\Phi(x){:}$ *is also a modal vector as, more generally, is* $\mathfrak{N}_1\mathfrak{N}_2 \ldots \mathfrak{N}_k\Phi(x){:}$ *provided merely that the* $\mathfrak{N}$*'s are the diagonal forms of any of the solutions of the equation* $\mu\mathcal{B}N{:} = N{:}$.

The modal numbers for the $\mathfrak{N}$*'s are among the roots of the equation*

$$det(\mu\mathcal{B} - \mathcal{I}) = 0.$$

---

[1] *Phil. Mag.*, **30**, 16 (1940); *Bull. Amer. Math. Soc.*, **48**, 153 (1942).

[2] A diagonal matrix of elements $a_1, a_2, \ldots, a_n$ is designated $[a_i)$.

[3] Fredholm, Ivar, *Acta Math.*, **27**, 371 (1903)

[4] Schmidt, Erhard, *Math. Annalen*, **63**, 461 (1907); in the notation of Stone, *Linear Transformations in Hilbert Space*, 1932, Schmidt's equations are written $\phi = T\psi$, $\psi = T^*\phi$, and T^* is called, following Schmidt, the "adjoint" of T.

[5] Birkhoff & Langer, *Proc. Amer. Acad. Arts Sci.*, **58**, 51–128 (1923).

[6] Kneser, *Integralgleichungen*, 2nd ed., 1922, pp. 199–212.

[7] Garbe, *Math. Annalen*, **76**, 538–542 (1914–1915).

[8] Mercer, *Proc. Roy. Soc. Lond.*, **A97**, 412 (1920).

[9] E. Schmidt, *Math. Annalen*, **63**, 446 (1907).

PROCEEDINGS
OF THE
NATIONAL ACADEMY OF SCIENCES

Volume 30 December 15, 1944 Number 12

Copyright 1944 by the National Academy of Sciences

THE MECHANISM OF BIOLUMINESCENCE

BY WILLIAM D. McELROY AND ROBERT BALLENTINE

PHYSIOLOGICAL LABORATORY, PRINCETON UNIVERSITY, AND DEPARTMENT OF ZOÖLOGY, COLUMBIA UNIVERSITY

Communicated November 3, 1944

Bioluminescent organisms offer unexcelled opportunities for an experimental attack on mechanisms of energy liberation and utilization in biological systems. Although the production of light by living organisms is of widespread occurrence, an actual demonstration of the essential components of the chemical systems involved has been made in members of only five of some twenty-one groups of luminous forms which have so far been investigated.[1] From these five an enzyme, luciferase, and its substrate, luciferin, have been separated. Although similar substances have not been demonstrated in other organisms, such as luminous bacteria, it is assumed that the basic reactions of all bioluminescence involves an enzymatic reaction, in the presence of oxygen, between a "luciferin" and its complementary enzyme, a "luciferase." Only from the crustacean, *Cypridina hilgendorfii*, have quantities of partially purified components of the luminescent reaction been isolated. The detailed study of the nature of the luminescent reaction itself has been limited to this isolated system. Consequently the recent quantitative formulation of inhibitor action in luminous bacteria[2] has rested upon the transfer of the properties of the demonstrated luciferin-luciferase system of *Cypridina* to the bacterial luminescent system. The general importance of this inhibitor theory is further reason for making every effort to understand more thoroughly the mechanism of light production.

Properties of the Cypridina Luminescent System.—Any mechanism for the bioluminescent reaction in *Cypridina* must be compatible with the following experimentally determined properties of this system.

(*a*) Luciferin undergoes two oxidative reactions. One of these is a reversible non-luminescent oxidation-reduction with an E_0' of $+ 0.26$ v. at pH 7.0.[3] The second is the following irreversible degradation of the side chain on the luciferin molecule:

$$\text{RCOCH}_2\text{OH} \rightarrow \text{RCOOH}^4. \qquad (1)$$

The latter is involved in the luminescent reaction proper and accounts for the over-all irreversibility in the isolated system.

(*b*) Both the light-producing and the non-luminescent reactions show the same series of changes in the absorption spectra. The first change observed is a shift in an absorption maximum at 430 mμ to 470 mμ. This is followed by a disappearance of the 470 mμ band. These changes occur over 100 times more rapidly in the luminescent than in the non-luminescent oxidation.[5]

(*c*) Luciferin contains a ketohydroxy side chain and a hydroquinone ring system,[4] probably in either the anthraquinone or naphthaquinone series.[5] Nitrogen, sulfur and halogen are absent from the molecule.[4]

(*d*) The emission spectrum of luminescence reveals peaks at 4750 Å. and 5600 Å.[6] These peaks correspond to energies of 59,400 and 50,000 cal.

(*e*) There is a release of phosphate during the luminescent reaction. This is correlated with the presence of acid-labile phosphate in the luciferin preparation.[7]

Except for point (*e*) the properties here listed have been known for some time. Nevertheless, as will now be shown, no over-all reaction scheme has yet been proposed as the explanation of the bioluminescent phenomenon which is consistent with all these properties. Following an analysis of the various schemes, we shall propose a hypothesis which not only is in agreement with all known properties, but which promises to be of service in devising new attacks on the problem.

Kinetic Interpretation of Bioluminescence.—Numerous studies on the kinetics of the bioluminescent reaction have led to the formulation of the following mechanism:[8]

$$\text{LH}_2 \text{ (luciferin)} + \text{A (luciferase)} \rightarrow \text{ALH}_2 \qquad (2)$$

$$\text{ALH}_2 + {}^1/{}_2\text{O}_2 \rightarrow \text{ALH}_2\text{O} \qquad (3)$$

$$\text{ALH}_2\text{O} \rightarrow \text{A}' \text{ (excited luciferase)} + \text{L (oxidized luciferin)} + \text{H}_2\text{O} \qquad (4)$$

$$\text{A}' \rightarrow \text{A} + h\nu \qquad (5)$$

The essential feature of this scheme consists in the combination of luciferase and luciferin (E_0' + 0.26, pH 7.0) followed by a direct transfer of two hydrogens to oxygen. In such a transfer the over-all free energy change is approximately 24,000 cal. This is insufficient to account for the energy release on luminescence as demanded by the emission spectrum data (point (*d*)). Inasmuch as this scheme is based upon the analysis of the light emitting step, it fails to consider the possible occurrence of "dark reactions" leading up to the final emission of light. Therefore it is not surprising that the points enumerated under (*a*), (*b*), (*c*) and (*e*) also remain unexplained by this scheme.

Johnson and Eyring[9] have attempted an elucidation of the nature of the luciferin molecule and the transformations undergone in the luminescent reaction. According to them "luciferin apparently contains both coenzyme (I or II) and a flavine prosthetic group, the former component providing a reductant, and the latter, after loose combination with its specific protein, comprising molecules excitable by oxidation." This hypothesis demands that the luciferin molecule possess a flavin component as an essential part. Contrary to Johnson and Eyring's observations, the chemical data do not support this assumption. First, it has been demonstrated that doubly purified luciferin does not contain nitrogen.[10] Second, independent of the nitrogen analysis, the flavin hypothesis is not in agreement with the spectroscopic data.[5] The three absorption peaks normally associated with known flavins are not to be found in the luciferin absorption spectrum. Further, it must be stressed that luciferin is initially in the reduced state (*a*), whereas the absorption bands of the flavin are characteristic of the oxidized molecule. Thus in the luminescent reaction the yellow color of luciferin disappears on oxidation rather than appearing (*b*) as would be expected if the substance were a flavin. Actually the absorption spectrum of luciferin shows a marked qualitative resemblance to certain naphtha- and anthraquinone derivatives.[5] Third, both Anderson and Korr place the redox potential near the hydroquinone-quinone system,[3] a value considerably more positive than those of known flavin systems. These data taken as a whole appear to invalidate the flavin hypothesis.[11]

An explanation of the irreversibility of the luminescent reaction was also proposed by the same authors.[9] According to their scheme "the excited molecules radiate and are not destroyed but others failing to radiate are destroyed by their absorbed energy." Regardless of the specific structure of the luciferin molecule, their hypothesis, nevertheless, implies that those molecules which emit light are reducible to luciferin again by simple reduction. But such a reversible oxidation is not experimentally demonstrable, and it has been shown that the luminescent reaction involves an irreversible degradation (Eq. 1).

Proposed "Dark Reactions" in the Luminescent Reaction.—The direct oxidation of two hydrogen atoms of glucose makes available 57,340 cal.;[9] however, from what is known of stepwise oxidation in biological systems it is doubtful that such a direct energy release is available for light emission. This essential energy requirement may well be provided by preceding "dark reactions" involving the degradation of the ketohydroxy side chain (Eq. 1) which is an essential step in the luminescent reaction. The pathway of this degradation may be postulated, in its simplest form, to occur by the following reaction steps (R represents the ring structure):

$$RCOCH_2OH \rightarrow RCOCHO + 2H \qquad (6)$$

$$RCOCHO + H_2O \rightarrow RCOCOOH + 2H \qquad (7)$$

$$RCOCOOH \rightarrow RCHO + CO_2 \qquad (8)$$

$$RCHO + H_2O \rightarrow RCOOH + 2H \qquad (9)$$

The removal of the hydrogen on the ring structure is probably the last of a series of reactions which make the light emitting molecule "energy rich." Calculations on the basis of analogous reactions which occur in cell oxidation[12] reveal that the energy release in the postulated series of reactions can be sufficient to provide the additional energy necessary for light emission.

This postulated mechanism is in agreement with available chemical information about the luciferin molecule (c) and provides an explanation of the irreversible nature of the luminescent reaction *in vitro* (*a*). At the present time there is insufficient knowledge of the structure of luciferin to make possible detailed interpretation of the absorption spectrum. However, the general nature of the luminescent and non-luminescent reactions as divulged by these spectral measurements does not contravert this mechanism. As has been pointed out earlier, the reversible oxidation of luciferin is presumably an oxidation-reduction of the ring structure.[8]

The absorption spectrum data suggest that the luminescent and non-luminescent reactions go through the same reaction chain. The question then arises as to how the energy from the side-chain oxidation is preserved to satisfy the energy requirements for light emission. From what is known of energy transformations in biological systems one might suspect that phosphate is concerned. If this is the case, the difference between the luminescent and the non-luminescent reactions would be that between phosphorolysis and hydrolysis. In the former the energy would be preserved as phosphate bond energy[13] whereas in the latter the energy would be lost as heat. The luminescent reaction here postulated is essentially a reversal of the reactions which have been proposed to take place in photosynthesis.[14] With such a mechanism, equations (3) to (5) of the kinetic scheme would be valid. The postulated series of reactions would merely replace equation (2).

Release of Phosphate on Luminescence.—Assuming the above mechanism, luciferin preparations might well contain measurable quantities of some "energy rich" phosphorylated molecules. As will now be shown, an acid-labile phosphate is readily detected in the luciferin preparation available to us.[15] Although it is difficult to rule out the possibility that the phosphate released was not actually present as a contaminant, the fact that phosphate is released (*vide infra*) during the luminescent reaction supports the conclusion that luciferin does indeed contain "energy rich" phosphorylated groups. The experimental data upon which these remarks are based may be presented as follows:

(1) Acid-labile phosphate. Doubly purified *Cypridina* luciferin, prepared by the method of Anderson,[16] was kindly supplied by Dr. A. M. Chase. One milligram of dry luciferin was dissolved in 15 ml. of 1 *N* hydrochloric acid. Aliquots were hydrolyzed at 100°C. for 7 and 30 minutes. Phosphate was determined by the method of Berenblum and Chain.[17] Inorganic phosphate initially present was 3.5, 3.7 γ/mg. luciferin preparation. This increased to 23.8, 24.5 γ/mg. at 7-minute hydrolysis. On 30-minute hydrolysis 22.4 γ/mg. was found. Hydrolysis under hydrogen for 7 minutes yielded 22.8 γ/mg.

(2) Phosphate release on luminescence. To 1.5 ml. of the above luciferin solution which had been brought to pH 7.4, 0.1 ml. of luciferase solution was added. The reaction was stopped by the addition of trichloroacetic when the major portion of light emission had taken place ($2^1/_2$ minutes). A control was run by adding trichloroacetic acid prior to the addition of the enzyme. The initial inorganic phosphate was found to be 9.4 γ/mg. luciferin preparation. After the luminescent reaction the inorganic phosphate had increased to 18.8 γ/mg. luciferin preparation.

---

[1] The numerous studies on bioluminescence are discussed by E. Newton Harvey, *Living Light*, Princeton University Press, Princeton, 1940.

[2] See Johnson, F. H., Eyring, H., and Kearns, W., *Arch. Biochem.*, **3**, 1–31 (1943).

[3] Korr, I. M., *Jour. Am. Chem. Soc.*, **58**, 1060 (1936); Anderson, R. S., *Jour. Cell. Comp. Physiol.*, **8**, 261–276 (1936); Harvey, E. N., *Ann. Rev. Biochem.*, **10**, 531–552 (1941).

[4] Chakravorty, P. N., and Ballentine, R., *Jour. Am. Chem. Soc.*, **63**, 2030 (1941).

[5] Chase, A. M., *Jour. Cell. Comp. Physiol.*, **15**, 159–172 (1940); *Jour. Biol. Chem.*, **150**, 433–445 (1943).

[6] Eymers, J. G., and van Schouwenburg, K. L., *Enzymol.*, **1**, 107–119 (1936).

[7] This paper.

[8] Chance, B., Harvey, E. N., Johnson, F., and Milliken, G., *Jour. Cell. Comp. Physiol.*, **15**, 195–215 (1940).

[9] Johnson, F., and Eyring, H., *Jour. Am. Chem. Soc.*, **66**, 848 (1944).

[10] The previously reported analysis for nitrogen[4] was of sufficient sensitivity to detect between 3–5 per cent nitrogen. This would mean that if the luciferin preparation was more than 20 per cent pure, a positive nitrogen test should have been obtained on the above flavin hypothesis.

[11] Data from other sources and unpublished observations on the correlation of flavin and luminescence in living systems would indicate that a flavin is probably concerned in the resynthesis of the luciferin molecule once it has been oxidized during luminescence. Doudoroff, M., *Enzymol.*, **5**, 239–243 (1938); Brooks, G., *Compt. Rend.*, **210**, 228–230 (1940); Ball, E. G., and Ramsdell, P. A., *Jour. Am. Chem. Soc.*, **66**, 1419–1420 (1944).

[12] Kalckar, H. M., *Chem. Rev.*, **28**, 71–178 (1941).

[13] Lipmann, F., *Advances in Enzymology*, Vol. 1, 99–162 (1941).

[14] Ruben, S., *Jour. Am. Chem. Soc.*, **65**, 297–281 (1943); Emerson, R. L., Stauffer, J. F., and Umbreit, W. W., *Am. Jour. Bot.*, **31**, 107–120 (1944).

[15] In testing for acid-labile phosphate in the purified preparations it was assumed that

the organisms were preserved under such conditions that some of the luciferin was partially oxidized and consequently remained "energy rich."

[16] Anderson, R. S., *Jour. Gen. Physiol.*, 19, 301–305 (1935).

[17] Berenblum, I., and Chain, E., *Biochem. Jour.*, 32, 295–298 (1938).

ON THE THEORY OF THE TENSION OF AN ELASTIC CYLINDER

By F. D. Murnaghan

Department of Mathematics, The Johns Hopkins University

Communicated October 20, 1944

In the accepted theory of the stretching of an elastic cylinder by a uniform traction in the direction of its axis, it is assumed that the cylinder which is granted to be elastically isotropic when the applied traction is zero, remains elastically isotropic after the traction is applied. We do not hold this assumption reasonable and we present here a theory of the traction of an elastic cylinder in which allowance is made for the lack of isotropy caused by the applied traction. Since we have very little definitive experimental knowledge of the effect of stress upon the elastic constants of a medium, our theory must be qualitative rather than quantitative. We take the point of view of our previous note[1] and so do not attempt to relate the energy of deformation to the (remote) unstressed state but, rather, consider merely the change of energy as we pass from any *stressed* position of the medium to a neighboring position in which the stress has been changed by an infinitesimal amount.

Taking our z-axis parallel to the axis of the cylinder, we assume that the stress is a uniform traction parallel to this axis so that the only non-vanishing component of the stress tensor is T_{zz} and we shall denote this simply by T. When T is changed to $T + \delta T$ any point (x, y, z) of the cylinder will be displaced to $(\xi, \eta, \zeta) = (x + \delta x, y + \delta y, z + \delta z)$. We follow the classical theory in assuming that δx is a function of x alone, that δy is a function of y alone and that δz is a function of z alone. This is only a rough approximation to what actually happens, since it implies that straight lines parallel to the axis of the cylinder are deformed into such and that plane cross-sections are deformed into such. As a consequence of this assumption the strain matrix e is diagonal with components $e_{xx} = \partial\delta x/\partial x$, $e_{yy} = \partial\delta y/\partial y$; $e_{zz} = \partial\delta z/\partial z$ (infinitesimals higher than the first order being neglected). We denote by ρ the density of the medium when under traction T and by ϕ the energy of deformation per unit mass when the traction is $T + \delta T$. In order to connect $T + \delta T$ and e we expand $\rho\phi$ as a function of the strain-components as far as terms of the second order. The first order terms in this expansion reduce to Te_{zz} since the stress, when

$e = 0$, is a uniform traction T parallel to the z-axis. The second order terms are complicated by the fact that the medium is not assumed to be, when under the traction T, elastically isotropic. It is, however, insensitive to rotations around the z-axis and so the second order terms are of the form

$$\tfrac{1}{2}\{(\lambda + 2\mu)I_1^2 - 4\mu I_2 + \alpha e_{zz}^2 + 4\beta e_{xx}e_{yy} + 2\gamma e_{zz}(e_{xx} + e_{yy})\}$$

where $I_1 = e_{xx} + e_{yy} + e_{zz}$, $I_2 = e_{yy}e_{zz} + e_{zz}e_{xx} + e_{xx}e_{yy}$. In writing down this expression we have availed ourselves of the simplification introduced by the fact that the strain matrix e is diagonal. The stress matrix, when the strain is e, is given[2] by $\left(1 + \frac{\delta\rho}{\rho}\right)P\,\frac{\partial(\rho\phi)}{\partial e}P^*$, where P is the matrix $\partial(\xi, \eta, \zeta)/\partial(x, y, z)$ and we must call attention to a point of some significance. When the initial state (from which e is measured) is unstressed, there are no linear terms in the expansion of $\rho\phi$ and, hence, no constant terms in the matrix $\partial(\rho\phi)/\partial e$. We may, therefore, since we are confining our attention to first order infinitesimals, replace $1 + \frac{\delta\rho}{\rho}$ by 1 and P by the unit matrix in the expression furnishing the stress matrix; this stress matrix is, accordingly, correctly given by the familiar expression $\partial(\rho\phi)/\partial e$. But this procedure is not valid when $\rho\phi$ contains the linear term Te_{zz}, for then the third diagonal element of the matrix $\partial(\rho\phi)/\partial e$ contains the constant term T.

Since we have assumed P to be diagonal, $\partial(\rho\varphi)/\partial e$ is diagonal and our stress matrix, when the strain matrix is e, appears in the form $\left(1 + \frac{\delta\rho}{\rho}\right)$ $(E + 2e)\,\frac{\partial}{\partial e}(\rho\phi) = (1 - I_1)(E + 2e)\frac{\partial}{\partial e}(\rho\phi)$, where E is the unit matrix. In order that the stress matrix be a uniform traction parallel to the z-axis we must have $e_{xx} = e_{yy} = -\sigma\, e_{zz}$ where $\sigma = (\lambda + \gamma)/2(\lambda + \mu + \beta)$ and then $\delta T = \left\{(\lambda + 2\mu + \alpha + T) - \frac{(\lambda + \gamma)(\lambda + \gamma - T)}{\lambda + \mu + \beta}\right\}e_{zz}$. Since $e_{zz} = \partial\delta z/\partial z$ is independent of z we have $e_{zz} = \delta z/z$ and so we obtain on integration

$$\log\frac{z}{L} = \int_0^T\{(\lambda + \mu + \beta)dT/[(\lambda + \mu + \beta)(\lambda + 2\mu + \alpha + T) - (\lambda + \gamma)(\lambda + \gamma - T)]$$

where L is the length of the cylinder when it is free from traction. In order to perform the integration we must know λ, μ, α, β, γ as functions of T. Let us assume, as a first approximation, that they are linear:

$$\lambda = \lambda_0 + lT,\ \mu = \mu_0 + mT,\ \alpha = aT,\ \beta = bT,\ \gamma = cT$$

(α, β, γ are zero when $T = 0$ since we assume the medium to be isotropic when free from stress). Then the denominator of the rational fraction which is to be integrated is the quadratic function $Q(T) = A + BT + CT^2$

where $A = \mu_0(3\lambda_0 + 2\mu_0)$; $B = \lambda_0(3m + a + b - 2c + 2) + \mu_0(3l + 4m + a + 2b + 1)$; $C = l(3m + a + b - 2c + 2) + (m + b)(2m + a + 1) - c(c - 1)$. If $Q(T)$ has a positive zero T_0 then $\log (z/L) \rightarrow \infty$ as $T \rightarrow T_0$; in other words the medium cannot support a traction greater than T_0.

If we disregard the effect of the traction T upon the elastic constants we have to set l, m, a, b, c all zero. We find

$$\frac{z}{L} = \left\{1 + (1 + 2\sigma)\frac{T}{E}\right\}^{\frac{1}{1 + 2\sigma}}$$

where $\sigma = \lambda/2(\lambda + \mu)$ is Poisson's ratio and $E = \mu\ (3\lambda + 2\mu)/(\lambda + \mu)$ is Young's modulus.

We remark in conclusion that an explanation of the yielding of the elastic cylinder under traction may be given without making the drastic hypothesis that the strain matrix is diagonal. The quadratic terms in the expansion of $\rho\phi$ are now

$$^1/_2\{(\lambda + 2\mu)I_1^2 - 4\mu I_2 + \alpha e_{zz}^2 + 4\beta(e_{xx}e_{yy} - e_{xy}e_{yx}) + 2\gamma e_{zz}(e_{xx} + e_{yy})\}$$

where $I_2 = e_{yy}e_{zz} + e_{zz}e_{xx} + e_{xx}e_{yy} - e_{yz}e_{zy} - e_{zx}e_{xz} - e_{xy}e_{yx}$, and the xy component of the stress is $2(\mu - \beta)e_{xy}$. In other words, the transverse rigidity of the cylinder is $\mu - \beta$ instead of μ (which would be its value if the cylinder were elastically isotropic). As $\beta \rightarrow \mu$ the medium becomes liquid-like in the sense of inability to support a (transverse) tangential stress. This may indicate why the cylinder suddenly *thins* at a certain point along its length.

[1] These PROCEEDINGS, 30, 244–247 (1944).

[2] Murnaghan, F. D., *Amer. Jour. Math.*, 59, 235–260 (1937).

THE SYNAPSIS OF THE SEX CHROMOSOMES OF DROSOPHILA MIRANDA IN RELATION TO THEIR DIRECTED SEGREGATION

BY R. H. MACKNIGHT AND KENNETH W. COOPER*

WM. G. KERCKHOFF LABORATORIES OF THE BIOLOGICAL SCIENCES, CALIFORNIA INSTITUTE OF TECHNOLOGY, AND DEPARTMENT OF BIOLOGY, PRINCETON UNIVERSITY

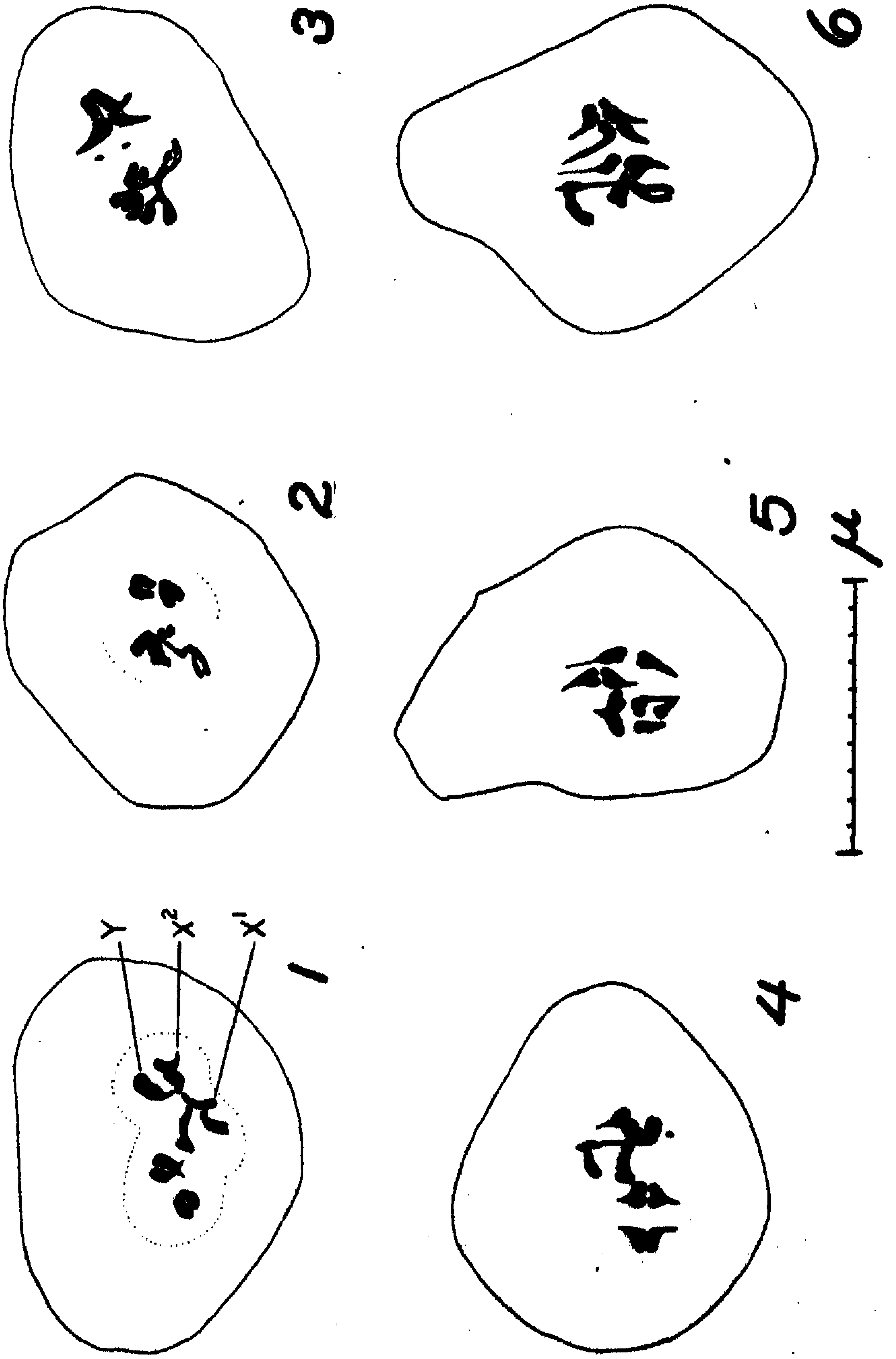
Communicated October 23, 1944

Dobzhansky[1] was the first to make a genetic and cytological study of *Drosophila miranda*. He concluded, among other things, that "the species has two distinct X chromosomes, X^1 and X^2. . . . The X^1 and Y chromosomes form a bivalent at meiosis. The X^2 chromosome remains unpaired . . . the X^2 passes at the first division to the same pole with the X^1. The nature of the mechanism insuring this determinate disjunction is unknown." MacKnight[2,3] further investigated the genetics and cytology of the species, and found evidence of partial homology between Y and X^2. Koller[4] studied spermatogenesis in a different geographic race of *D. miranda* as well as in the original strain, and in hybrids between these. Referring to certain hybrid material in which he thought abnormal chromosome behavior had been induced by high temperature, Koller concludes that "In a few instances chromosome configurations suggesting the association of the X^1Y bivalent with X^2 were observed in the hybrid." He figures several such configurations. Referring, however, to the pure races and to the hybrids when not heat treated, he states, "The position of the single X^2 varies within the spindle... (it) lies off the equator, but it has been found not infrequently amongst the autosomal bivalents."

A cytological study of the first spermatocyte division has been made, using males of a stock descended from the Olympic-1 strain collected by Dobzhansky, as well as males of the strains known as Whitney-60 and Big-Basin-2.[5] It has not been possible to confirm Dobzhansky's conclusion that X^2 remains unpaired, nor Koller's that the position of the single X^2 varies within the spindle. Rather a configuration consisting of two rod-bivalents, a dot-bivalent, and a sex-chromosome multivalent has been found in each of eighteen cells in Olympic-1 males, while no cell showing X^2 unsynapsed has been encountered (see Fig. 3).

The results are essentially the same for males of Whitney-60 and Big-Basin-2. At late diakinesis and prometaphase less than 5 per cent of the figures of cysts favorable for study are undecipherable or questionable, the remainder showing unequivocal sex chromosome trivalents (Figs. 1 and 2). At first metaphase roughly 15 per cent of the cells cannot be deciphered, but a free X^2 was found only eight times in a total of 239 completely analyzable meiotic complements, the remaining 231 configurations giving unequivocal sex chromosome trivalents (Fig. 4). In each of the eight exceptional cases it is probable that X^2 had precociously disjoined from an antecedent X^1X^2Y trivalent.[6] Furthermore early anaphase (Figs. 5 and 6) consistently show X^1 and X^2 alike disjoining from Y. In only one of 46 cases of mid or late first anaphase was X^2 found at the equator as described by Dobzhansky; 44 of the remaining figures showed X^1 and X^2 close to the pole opposite that toward which Y progressed, while in one case X^2 could not be identified.

Although the precise nature of the association between X^1, Y and X^2 has not been determined, the observation of synapsis between them makes it unnecessary to believe that *D. miranda* is one of those few organisms which accomplish the directed segregation of chromosomes not in contact at metaphase. Directed segregation of three or more chromosomes which do come in contact at metaphase is not uncommon, occurring in *Oenothera, Datura, Humulus, Tenodera* and numerous other organisms. Listing *D.*



FIGURES 1–6

(For explanation, see opposite page)

miranda with these does not explain the phenomenon, which remains one of the most unaccountable features of chromosome behavior. The tentative explanation offered by MacKnight[7] for some of these organisms is evidently not applicable to *D. miranda.*

Grateful acknowledgment is made to Professor Dobzhansky for suggesting this species as an object of investigation to R. H. MacKnight, and for making the Olympic strain available. To Professor A. H. Sturtevant we wish to express our joint appreciation for encouragement and advice.

---

* John Simon Guggenheim Memorial Foundation Fellow.

[1] Dobzhansky, Th., *Genetics*, **20**, 377–391 (1935); see also *Jour. Genetics*, **34**, 135–151 (1937).

[2] MacKnight, R. H., *Genetics*, **23**, 158 (1938).

[3] MacKnight, R. H., *Ibid.*, **24**, 180–201 (1939).

[4] Koller, P. C., *Jour. Genetics*, **38**, 477–492 (1939).

[5] Observations on the Olympic-1 strain were carried out solely by R. H. MacKnight, whereas Whitney-60 (derived from the same locality as Koller's "Whitney race") and Big-Basin-2 were independently studied by K. W. Cooper. The latter strains were kindly made available for investigation by Prof. A. H. Sturtevant.

[6] Unpublished observations of Dr. Hans Ris of Johns Hopkins University lead to an identical conclusion, namely, a sex chromosome trivalent is normally formed in the spermatogenesis of *D. miranda.*

[7] MacKnight, R. H., *Collecting Net*, **15**, 171 (1940).

EXPLANATION OF FIGURES

Stages of the first meiotic division in the male of *Drosophila miranda* Dobzhansky showing the autosomal bivalents and the sex chromosome trivalent.

Figure 1, diakinesis. Y is a thick, V-shaped chromosome to one arm of which X^2 is synapsed. X^1 is a slender, V-shaped chromosome synapsed interstitially with Y and in the vicinity of the kinetochore of Y. Figure 2, prometaphase. The disposition of the elements of the sex chromosome trivalent is roughly the mirror-image of that of figure 1. Figure 3, metaphase showing two rod-bivalents (superimposed), a dot-bivalent and a multivalent. Figure 4, metaphase. The disposition of the elements of the sex chromosome trivalent mirror-images that of figure 1; one element of the dot-bivalent appears below the lower arm of Y. Figure 5, early anaphase. The sex chromosome trivalent is to the left and shows the Y (above) disjoining from X^1 (to right, below Y) and X^2 (to left, below Y). Figure 6, early anaphase. The sex chromosome trivalent is to the left, with X^1 (ultimate, to left) and X^2 (penultimate, to left) above, Y below.

Figures 1, 4 and 6 are of Whitney-60. Figures 2 and 5 are of Big Basin-2. Figure 3 is of Olympic-1. The scale below figure 5 represents 10 micra and accompanies figures 1–2 and 4–6.

VISIBLE EXPRESSION OF CYTOPLASMIC PATTERN IN THE DIFFERENTIATION OF XYLEM STRANDS

By Edmund W. Sinnott and Robert Bloch

Osborn Botanical Laboratory, Yale University

Communicated October 28, 1944

In the differentiation of tissues during plant development, continuous series of cells undergo similar changes which distinguish them from adjoining cell groups. Thus at the terminal growing region of a stem the primary vascular bundles become differentiated in the midst of parenchymatous tissue. Such changes occur along precise pathways which form a definite pattern, but what determines the course of these pathways is unknown. We may speak of a "field" in which differentiation occurs, though this is little more than a descriptive term. Presumably the particular cells which are to form a differentiated cell group first undergo chemical or physical changes which are the precursors of visible morphological alterations, but in most cases these are difficult or impossible to determine. The cells of plants, however, possess an important advantage in this regard, for in them, except in very small ones, the cytoplasm typically is dispersed in a thin layer just inside a relatively heavy cellulose wall, and it is sometimes possible to observe in this thin layer some of these preliminary changes and thus to study differentiation in its very early stages.

In the regions of normal differentiation cells at the meristem usually are small and rather densely cytoplasmic so that these changes are relatively difficult to observe. It is often possible, however, to induce specific types of differentiation in cells which are large, vacuolate and nearly mature, and thus to see much more clearly the early steps in the process. Thus by severing one or more already differentiated vascular bundles in the developing stem of a herbaceous plant, strands can be induced to develop in the parenchymatous tissues of the pith which link up the severed bundle with others and thus regenerate a functioning conductive system. Regeneration of this sort has been described by a number of investigators.[1, 2] Such differentiation in its fundamentals is essentially like that which occurs in normal development, but for a study of the process it has the important advantage that it is not complicated by growth of the tissues, since growth here has largely ceased before differentiation takes place.

The writers have studied vascular regeneration in several herbaceous plants, especially *Coleus hybridus*. Where pieces of stem were cut out laterally in still enlarging internodes, new strands were observed to develop readily in various places. Sometimes these linked up the severed ends of a bundle. Sometimes they connected one of these ends with another intact strand adjacent to it. Occasionally they were formed across the pith to a bundle on the other side (Fig. 1).

In the development of such a strand, there are several very early indications of the course along which differentiation will occur. Among the essentially isodiametric parenchyma cells in the path of the new strand, divisions often take place, and these are predominantly parallel to the direction of the strand (Fig. 2) and ultimately lead to the formation of elongate vascular initials. Each division, as in all vacuolate cells, is preceded by the formation of a plate of cytoplasm (the phragmosome[3]) in the position of the future wall, so that in these cells the orientation of the cytoplasm must have been altered to conform to the axis of the new strand.

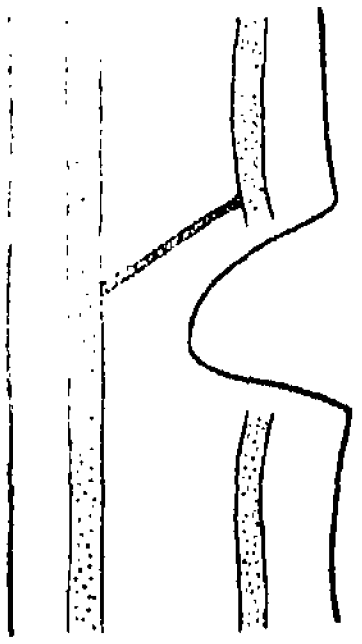


FIGURE 1

Diagrammatic longitudinal section through internode of *Coleus* eight days after wounding, showing normal vascular strands (dotted) and regenerated connection between them across the pith above the wound gap.

Cytoplasmic changes occur in all the cells of the future bundle, but a clearer picture of them is to be seen in the development of the characteristic wall thickenings in those cells which are to form the xylem. In most such cells the lignified secondary wall is formed in rings or spirals around the cell, or in a reticulate system with long, narrow pits. Many of the cells are vessels with pores or pore-like areas in their walls. In the elongated cells cut out of the basic parenchyma, these thickenings are much like the ones in typical protoxylem or primary xylem, and the rings or spirals run at right angles to the long axis of the cell. In those cases, however, where the large, isodiametric parenchyma cells differentiate directly into xylem, the orientation of these thickenings shows an important relation to the course of the strand. Figure 2 is of such a regenerative strand in which the xylem cells retain essentially the shape of the original parenchyma. It is evident that even in the cells which are approximately square in section, the direction of these lignified bands tends to be at right angles to the course of the strand. Furthermore, the bands in one cell are directly opposite those in adjacent cells so that series of bands, approximately equidistant, form sweeping, contour-like curves passing across the whole cell group (Fig. 3).

In this system of bands, and often appearing as distortions or "eddies" in them, are pores. Their position is also related to the course of the differentiating strand for they do not develop at the originally apical and basal ends of these cells but in such a position that the line of pores or poreike areas is parallel to the axis of the strand (Fig. 2).

The significance of these facts becomes clear when the origin of the lignified bands themselves is understood. Many years ago Crüger,[4] Dippel[5] and other observers who studied the genesis of such thickenings in the secondary walls of plant cells saw that along the course of what was to be a lignified thickening there is a band of densely granular cytoplasm. In some cases this could be seen to be streaming. Such bands were still visible in the contracted protoplast of plasmolyzed cells in which the walls had not yet developed thickenings. These facts indicate that there is a definite configuration of the cytoplasm which precedes but is similar to the configuration of lignified thickenings, and which is presumably related to their development. These important observations seem largely to have been overlooked by British and American cytologists and histologists.

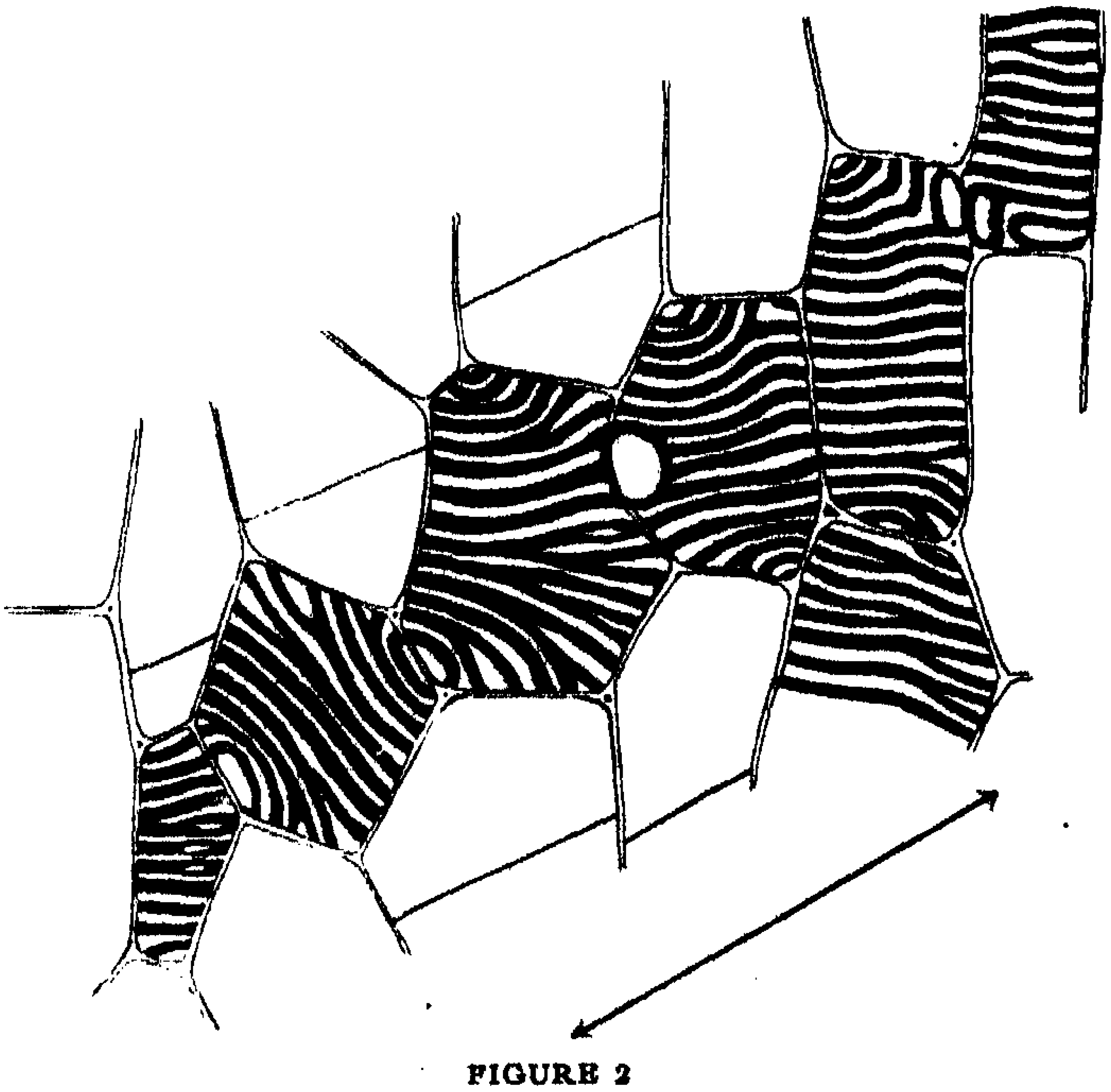


FIGURE 2

Enlarged portion of a regenerating xylem strand as shown in figure 1, passing obliquely through pith parenchyma from edge of normal strand (right), showing pattern of lignified bands in walls and position of pore areas. Several parenchyma cells have divided, the new walls being parallel to course of strand (shown by arrow). Eight days after wounding. Camera lucida drawing, × 400.

The present writers have frequently confirmed their accuracy. Such bands of granular cytoplasm are often difficult to see in the small provascular cells in normal development, but are frequently demonstrable in the xylem of regenerating vascular strands. In many cases they constitute the first visible indication of what cells are to comprise the new strand and what its course is to be. Their later embodiment in the lignified cellulose pattern of the cell walls thusr povides a picture, so to speak, of the field in which the strand has differentiated.

Evidently the first visible change in the cells which are to form a new strand is a reorientation of the cytoplasm in them. This presumably occurs in all cells, but can be demonstrated particularly well in those of the xylem, where the cytoplasm has a characteristic pattern of its own. This pattern, as described above, first expresses itself in a cytoplasmic system the polarity of which, as shown by its orientation, no longer conforms to that of the original parenchyma cells but to the direction of the new strand. Whatever forces determine the course and differentiation of this strand seem to do so by acting upon the living material of the cells out of which it is formed. The nature of these forces is yet far from clear, but the rhythmic cytoplasmic contour patterns which sweep across whole groups of cells perhaps have more than a superficial resemblance to certain configurations observed in colloidal systems, such as Liesegang rings,[6] or to the lines of force in an electrical field. The character of such configurations may therefore provide important evidence in an analysis of the factors which control differentiation.

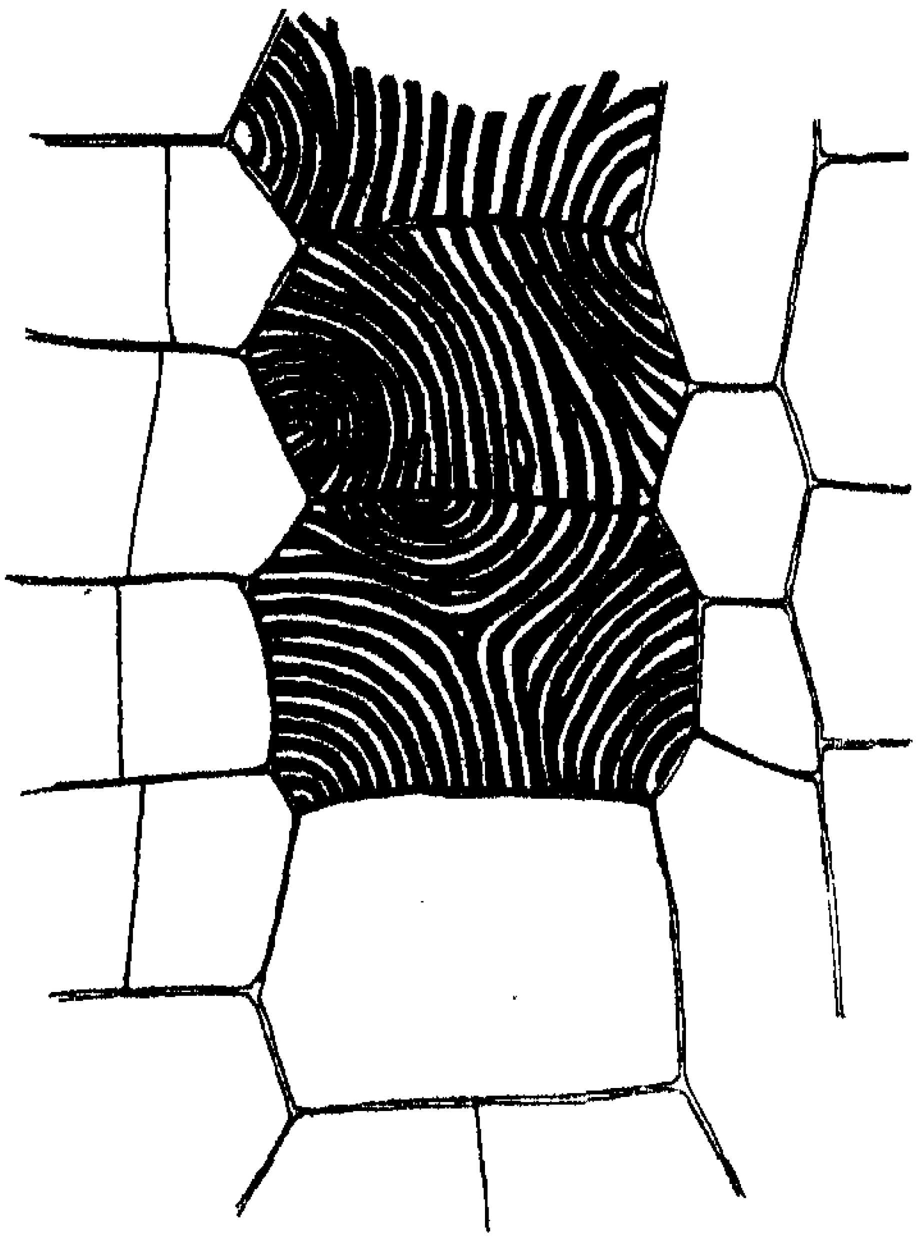


FIGURE 3

Portion of another regenerating strand showing three individual xylem cells the lignified bands of which form a continuous pattern. Camera lucida drawing, × 400.

Since such a pattern or field transgresses cell boundaries and involves many cells, an analysis of it must give particular attention to the relations which exist betweeen a cell and its neighbors. This is a problem which has attracted little attention from plant histologists and from cytologists generally, who have been primarily concerned with the character and establishment of types of individual cells. Plant cells with their sculptured walls provide good opportunities to study such relationships. Thus in ringed or spiral elements of the protoxylem, observation readily establishes the fact that the lignified thickenings in one cell are directly opposite those in the next, as has been shown in the present paper. The fact that pit openings are opposite each other is the necessary result of such relationships. More complex examples of intercellular wall patterns extending in continuous series across many cells are those of the supporting network (*réseau de soutien*) in the cortex of orchid air roots, the wall thickenings of the anther endothecium in many plants, and the continuous series of division walls in such tissues as wound cork.[7] In these cases and others it is clear that the cytoplasmic changes which occur in one cell during the process of differentiation are not isolated events but are part of a connected series of changes which involve a whole group of interrelated cells. A study of these relationships forms an important part of the developmental analysis of histological patterns.

Summary.—In the regeneration of a vascular strand in the pith, the cytoplasm of cells originally parenchyma but now destined to become xylem visibly assumes a new configuration which conforms to the course of the regenerating strand. These changes later express themselves in the development of characteristic lignified bands which thus make permanently visible the cytoplasmic changes in the field of the regenerating strand.

[1] Simon, S., *Ber. deutsch. bot. Ges.*, **26**, 364–396 (1908).

[2] Freundlich, H. F., *Jahrb. Wiss. Bot.*, **46**, 137–206 (1909).

[3] Sinnott, E. W., and Bloch, R., *Amer. Jour. Bot.*, **28**, 225–232 (1941).

[4] Crüger, H., *Bot. Zeit.*, **13**, 601–613, 617–629 (1855).

[5] Dippel, L., *Abhandl. Naturforsch. Ges. Halle*, **10**, 53–68 (1867).

[6] Küster, E., *Über Zonenbildung in kolloidalen Medien*, 2. Aufl., Jena (1931).

[7] Sinnott, E. W., and Bloch, R., *Amer. Jour. Bot.*, **28**, 607–617 (1941).

A GROWTH-DELAYING EFFECT OF ULTRAVIOLET RADIATION ON BACTERIAL VIRUSES

BY S. E. LURIA

BACTERIOLOGICAL LABORATORIES, INDIANA UNIVERSITY*

Communicated October 20, 1944

Ultraviolet and ionizing radiations are known to inactivate bacterial viruses (bacteriophages).[1, 2] The inactivated virus no longer multiplies on sensitive bacteria or causes lysis. In the case of ultraviolet radiation, it has been found in some instances that the inactivated virus still retains the ability to be adsorbed by sensitive bacteria, and to interfere both with their viability and with their ability to support growth of active virus.[3] This residual property of irradiated virus is itself destroyed by higher doses of radiation.

Radiation is supposed to produce mutations in viruses.[4] In the course of attempts to increase by irradiation the rate of occurrence of mutations affecting the host range of bacterial viruses,[5] we encountered a new type of physiological, non-hereditary effect of ultraviolet radiation, namely, a growth-delaying effect, which will be described in the present communication.

Bacterial viruses grow by infecting living bacterial cells of sensitive strains and multiplying inside the cells. After a fairly constant interval of time, the cells are lysed and liberate a large number of virus particles.[6] We can describe the process in terms of a series of parameters: rate of virus adsorption by bacteria; "constant period," i.e., minimum interval between infection and lysis; "rise period," i.e., spread in the values of this interval for individual cells; "burst size," i.e., average yield of virus particles per lysed cell. These parameters are easily reproducible under given conditions for each virus-host system. Methods for their determination have been described in detail.[7]

The radiation effect with which we are dealing was discovered when we studied by these methods the growth of the fraction of irradiated virus that had survived ultraviolet irradiation. Filtrates of various bacterial viruses, all active on a common host, *Escherichia coli B*, were irradiated in thin layers (average about 0.02 mm.) with ultraviolet light containing about 80% of 2537 Å. radiation. The virus surviving various times of irradiation was then titrated and, within a few minutes, growth experiments of this surviving virus were performed. Our conditions of irradiation, although not such as to eliminate all screening effects by components of the medium, were, however, accurate enough to give reproducible amounts of inactivation.

The three viruses studied, α, γ and $T7$, are known to be unrelated on the basis of serological and cross-resistance tests. Viruses α and *T7* consist of small particles, virus γ of large particles, as shown by electron micrography and x-ray sensitivity tests.[8, 9] Upon ultraviolet irradiation, virus γ proved the most sensitive, virus *T7* more resistant, virus α very resistant. For the same percentage of inactivation, the relative doses for the three viruses, measured in time of exposure, were in the approximate ratios of 1:4:10.

In the growth experiments using the virus surviving irradiation, adsorption by sensitive bacteria proved normal. When the constant periods were measured, however, they were found to be longer than for normal virus. The rise period was also longer than normal, whereas the burst size was normal. The values for the constant period are shown in table 1.

TABLE 1

	TIME OF IRRADIATION, MIN.	RESIDUAL ACTIVITY, %	CONSTANT PERIOD OF GROWTH ON *Escherichia coli B* (IN NUTRIENT BROTH AT 37°C.), MIN.
Virus α	0	100	13
	1	20	18
	3	2	25
	6	0.3	31
Virus γ	0	100	21
	0.5	1	24
Virus *T7*	0	100	13
	1	7	23
	5	0.2	28

Rise period and burst size were only determined in a few cases, because it is difficult to run complete growth experiments with low initial virus titers. We had to avoid the use of concentrated virus suspensions, in order to exclude the possibility of interference effects from the inactive virus particles: all experiments were done with bacteria in excess of the total amount of virus (active + inactive). The increase in rise period, however, was clearly shown in all cases by changes in the slope of the growth curve of virus.

As a whole, the effect appears to consist of a delay in virus liberation, without change in virus yield per bacterium. Not only is the minimum time for lysis longer, but the intervals between infection and lysis for individual bacteria are also more variable, as shown by the increase of the rise period. Both constant period and rise period are progressively more affected, for each virus, by increasing doses of radiation.

Similar results were obtained if the irradiated virus was diluted and stored at 5°C. for one day between irradiation and the growth experiment. The effect proved, however, strictly non-hereditary: the offspring of the irradiated virus, when tested in its turn, grew on sensitive bacteria in the same way as normal virus.

It is seen in table 1 that the growth-delaying effect of radiation, while increasing with the dose of radiation, does not parallel the inactivation effect on the various viruses: for comparable amounts of inactivation, growth delay is pronounced for the slowly inactivated viruses α and $T7$, barely evident for the very sensitive virus γ. It appears that the growth-delaying effect is more closely related, for various viruses, to the total time of irradiation, that is, to the number of quanta absorbed per unit of volume. Since the delay effect must be brought about by the action of quanta which have not acted lethally, and since it is a typically progressive effect, increasing in intensity with increasing doses of radiation, one is led to conclude that it results from a cumulative effect of ultraviolet quanta on the body of the virus particles.

Upon further irradiation, these particles would then become inactivated. Inactivation is likely to be caused, not by the cumulative effect of a larger number of quanta, but by a single "effective hit," as is true in the case of ionizing radiation, and as seems suggested by Gates's results with ultraviolet light.[1] A discussion of the mechanism of inactivation is, however, beyond the scope of the present paper.

The discovery of a physiological effect of non-lethal ultraviolet quanta on the particles of bacterial viruses raised an interesting question as to the occurrence of a similar effect for x-rays. It has been conclusively proved[8, 9] that in the case of x-rays nearly each "hit," that is, each elementary act of absorption of radiation in a particle of a bacterial virus, is effective in producing inactivation. A comparison of the "sensitive volumes" with the actual volumes of the particles shows that the ratio of hits to effective hits is not larger than 3:1 or 4:1, and is very likely close to 1:1. Under such conditions, we should hardly expect any non-lethal effect of radiation energy absorbed by the particles, comparable to the effect described above for ultraviolet light.

This expectation was proved to be correct in a series of experiments in which we exposed viruses α and γ to doses of x-rays (200 kv.) sufficient to reduce their titers to 0.1–5 per cent, and then studied the growth of the surviving virus. In all cases the growth was completely normal: constant period, rise period and burst size were like those of non-irradiated controls. It is clear that no delaying effect on growth is produced by x-rays, as expected in view of the high yield of inactivation by each act of absorption.

The effect of ultraviolet light described in this paper is an example of a non-lethal, non-genetic alteration of virus particles produced by radiation. The change involved affects the time interval between adsorption of a virus particle by a sensitive host and liberation of new virus. It is impossible to decide at present which of the various steps involved in the process leading to virus liberation is affected, whether it is the speed of penetration into the bacterial cell, or the rate of multiplication inside the cell, or the reactions responsible for cell lysis. A delay in any one or more of these steps would account for the result. It is worth recalling that non-lethal, growth-delaying effects of ultraviolet and ionizing radiation are of familiar occurrence in the case of bacteria, yeasts, protozoa and cells of higher organisms. The growth-delaying effect here described may well be the expression of the same phenomenon in bacterial viruses.

The actual nature of the structural change causing growth delay is also open to speculation. Since bacterial viruses seem to be composed of nucleoproteins [10, 11] a study of the effectiveness of ultraviolet light of different wave-lengths, differently absorbed by specific components of nucleoproteins, might throw some light on this question.

In conclusion, it is seen that virus particles, in spite of their small size and probably simple composition, can undergo, under the influence of ultraviolet radiation, structural changes of various kinds. Of these, some affect the reproducing capacity of the virus; others even destroy the ability of the virus to be adsorbed by the host cells and to interfere with the growth of other virus particles; other changes may hereditarily affect some of the properties of the virus particles; finally, changes of the kind here described alter only the rate of some reactions involved in virus reproduction, in a strictly individual, non-hereditary manner.

The growth-delaying effect of radiation, detected for bacterial viruses, may well occur with other groups of viruses. In most of them, however, such an effect would be difficult to detect, since only for bacterial viruses can we study an isolated step of growth of a given group of virus particles. Radiation effects of this type may be of some relevance in the study of virus attenuation, partial neutralization, vaccine production and similar problems.

Summary.—Ultraviolet radiation, besides inactivating bacterial viruses, produces a delay in the growth of the surviving virus particles on a sensitive host; the delay increases with increasing doses of radiation. This effect is non-hereditary, and differently pronounced for different viruses. It appears to be due to the cumulative effect of the quanta absorbed by the virus particles before their inactivation. No similar effect was found on virus particles which had survived x-ray irradiation. This is explained by the fact that, for ionizing radiation, nearly each act of absorption by a virus particle is effective in producing inactivation: the surviving particles are likely not to have absorbed any radiation at all.

[*] Part of the experiments described in this paper were done in the summer of 1944 in the Department of Genetics of the Carnegie Institution of Washington, Cold Spring Harbor, New York.

[1] Gates, F. L., *Jour. Exptl. Med.*, **60,** 179 (1934).

[2] Wollman, E., and Lacassagne, A., *Ann. Inst. Past.*, **64,** 5 (1940).

[3] Luria, S. E., and Delbrück, M., *Arch. Biochem.*, **1,** 207 (1942).

[4] Gowen, J. W., *Cold Spring Harbor Symp.*, **9,** 187 (1941).

[5] Luria, S. E., *Genetics*, in press.

[6] Delbrück, M., *Advances in Enzymology*, Vol. 2, 1 (1942).

[7] Delbrück, M., and Luria, S. E., *Arch. Biochem.*, **1**, 111 (1942).

[8] Luria, S. E., and Anderson, T. F., these PROCEEDINGS, **28**, 127 (1942).

[9] Luria, S. E., and Exner, F. M., *Ibid.*, **27**, 370 (1941).

[10] Schlesinger, M., *Biochem. Z.*, **264**, 6 (1933).

[11] Northrop, J. H., *Jour. Gen. Physiol.*, **21**, 335 (1938).

INCIDENCE OF METABOLIC CHANGES AMONG VIRUS-RESISTANT MUTANTS OF A BACTERIAL STRAIN

BY E. H. ANDERSON*

DEPARTMENT OF BIOLOGY, VANDERBILT UNIVERSITY

Communicated October 25, 1944

It has long been known that a virus-sensitive strain of bacteria which is subjected to the action of a virus (bacteriophage) will give rise to a secondary growth of organisms which will be specifically resistant to the virus causing the lysis of the parent strain.[1] However, it was not until the investigations of Luria and Delbrück[2] that the mechanism for the origin of virus-resistant bacteria was conclusively established. Their analysis of the distribution of the numbers of virus-resistant bacteria in a series of similar cultures of a virus-sensitive strain of *Escherichia coli* validated the hypothesis that the resistant bacteria arise by mutation of the sensitive cells independently of the action of the virus. The virus, therefore, in destroying the sensitive cells serves as a selective rather than as a causative agent.

In general, resistant mutants do not exhibit correlated colony, morphological or metabolic characteristics which differ from those of the parent strain. In the majority of cases the only criterion by which a mutant may be distinguished from the original strain is by its altered reaction to the virus or viruses in whose presence it developed. In the case of the sensitive strain the virus particles are readily adsorbed onto the bacterial cells which, after a characteristic latent period, undergo lysis. The resistant mutant, on the other hand, is not lysed nor will it show any affinity for the viruses for which it is resistant. This resistance is a stable heritable character as is evidenced by the fact that the mutant will retain its specific resistance even when carried through many subcultures in the complete absence of viruses.

The demonstration that bacterial resistance to a virus may be traced to a specific genetic change[2] may be assumed to indicate that the mutant must differ in some profound manner from the strain from which it arose. Such a mutation may involve solely an alteration in the structures or mechanisms responsible for the response of the cell to a specific virus, or it may embody a change of a more fundamental nature and involve a correlated change in the metabolic systems of the cells as well. A mutation of the latter type would constitute a biochemical mutation and presumably could be detected by a comparison of the growth requirements and the metabolic activities of the mutants with those of the parent strain.

The purpose of the present study was to isolate a number of mutant strains of *Escherichia coli* whose sensitivities to various viruses differ from each other and from those of the parent strain and to determine if they possess any differences in their biochemical activities which may be correlated with their differences in response to the infecting agents. Indications that a mutation from virus sensitivity to virus resistance may be coupled with a biochemical mutation would be of great value to any attempt to obtain a biochemical interpretation of the relation of bacterial virus growth to the metabolic activities of the host cell.

Methods and Results.—The organism used as the parent strain from which mutants for this study were obtained was the same B strain of *Escherichia coli* as that used previously by Delbrück and Luria.[3] Mutants were obtained by subjecting the B strain to the action of three virus strains, all of which were active on this host. The viruses used were the strains α and γ described by Delbrück and Luria[3, 4] and a new strain, δ, which was isolated in this laboratory and which has been determined to be identical with strain *T7* of Demerec and Fano.[5]

The method of obtaining the resistant mutants consisted of setting up a series of test tubes each containing 0.15 ml. of a very dilute suspension of bacteria in Difco nutrient broth to which had been added 0.5 per cent sodium chloride. After an 18- to 24-hour period of incubation at 37°C., at which time the bacterial concentration had reached saturation, the total contents of each tube were removed by pipetting and spread over the central surface of nutrient agar plates which had been previously prepared by spreading their entire surface with an excess of a single strain of high titer virus stock. In this manner all bacteria were surrounded by large numbers of virus particles which destroyed all the sensitive cells. The few colonies of resistant bacteria which developed on these plates were allowed to grow for about 48 hours. Representative colonies were then picked from each plate and the strains isolated were freed from all virus particles by serially streaking out on at least two additional nutrient agar plates. Each mutant thus isolated in pure culture was tested for its resistance to the virus in whose presence it was developed and only those showing complete resistance were retained.

In this manner primary mutants were obtained from the parent strain which were resistant to any one of the three viruses and sensitive to the other two. By utilizing these single mutants in the above procedure double mutants were obtained which were resistant to any two viruses and sensitive to the remaining one. Mutant strains of B were also obtained which were completely resistant to α, γ and δ.

Fifty-seven virus-resistant mutants, arising from independent mutations, have been isolated in pure culture. These represent seven different groups on the basis of their reactions to the three viruses. While many representatives were isolated because of colony variation, a number of the strains may be the result of independent duplicate mutations. Others were selected because of the difference in the order in which their resistance to two viruses developed.

Of 54 strains tested, all agree in their morphological characteristics, the cells appearing as gram-negative, non-motile rods. Three types of surface colony formations are distinguishable in the pure cultures of the virus-resistant organisms that have been isolated. These have been classified as large, which correspond in size and shape to those of the parent organism; small, which are approximately one-half the size but similar in shape and appearance to the first type; and tiny, which are very small and require a longer time to develop than do the other two types. Although the small type is found more frequently among the mutants showing resistance to the α virus there does not appear to be any correlated relationship between type of resistance and colony formation.

The method developed by Beadle and Tatum[6] for detecting biochemical mutations in their studies on the genetic control of biochemical reactions in the ascomycete *Neurospora* seemed ideally suited for use in determining whether or not the virus-resistant strains of coli require factors for growth other than those required by the original B strain. The method is based on the assumption that any mutation which blocks the synthesis of an indispensable diffusible substance normally synthesized by the parent organism would be lethal in an environment which fails to supply the product of the blocked synthesis. Under environmental conditions in which the product is readily available in a utilizable form the metabolic deficiency would be compensated for and the organism would survive despite the biochemical mutation. The virus-resistant mutants were developed and carried on nutrient agar medium containing a large number of normally synthesized constituents of cells and thus would be capable of growth even though a mutation had taken place which blocked the synthesis of any of the products contained in this "complete" medium. A lost function of this nature could be detected, however, on transfer from the complete medium to one supplying only the minimum requirements of the parent strain. The inability of a mutant to develop in a "minimal" medium would indicate that it had lost the power of synthesizing some indispensable substance present in the complete medium and normally produced in the virus-sensitive strain.

A minimal medium containing inorganic salts, asparagine and dextrose, buffered at pH 7.0, has been found adequate to support the growth of the B strain. Although the parent organism grew more slowly in this medium, in that under identical conditions of temperature and aeration it had a generation time of 40 minutes as compared with 19 minutes in broth, it was able to reach the same or higher concentrations as in broth. When all 57 mutants were transferred to the minimal medium some of them grew as abundantly as did the B strain. Others, however, grew more slowly and *28 strains were unable to grow at all.* The addition of minute amounts of yeast extract (0.005 per cent Difco yeast extract) to the minimal medium enabled all strains to show good growth.

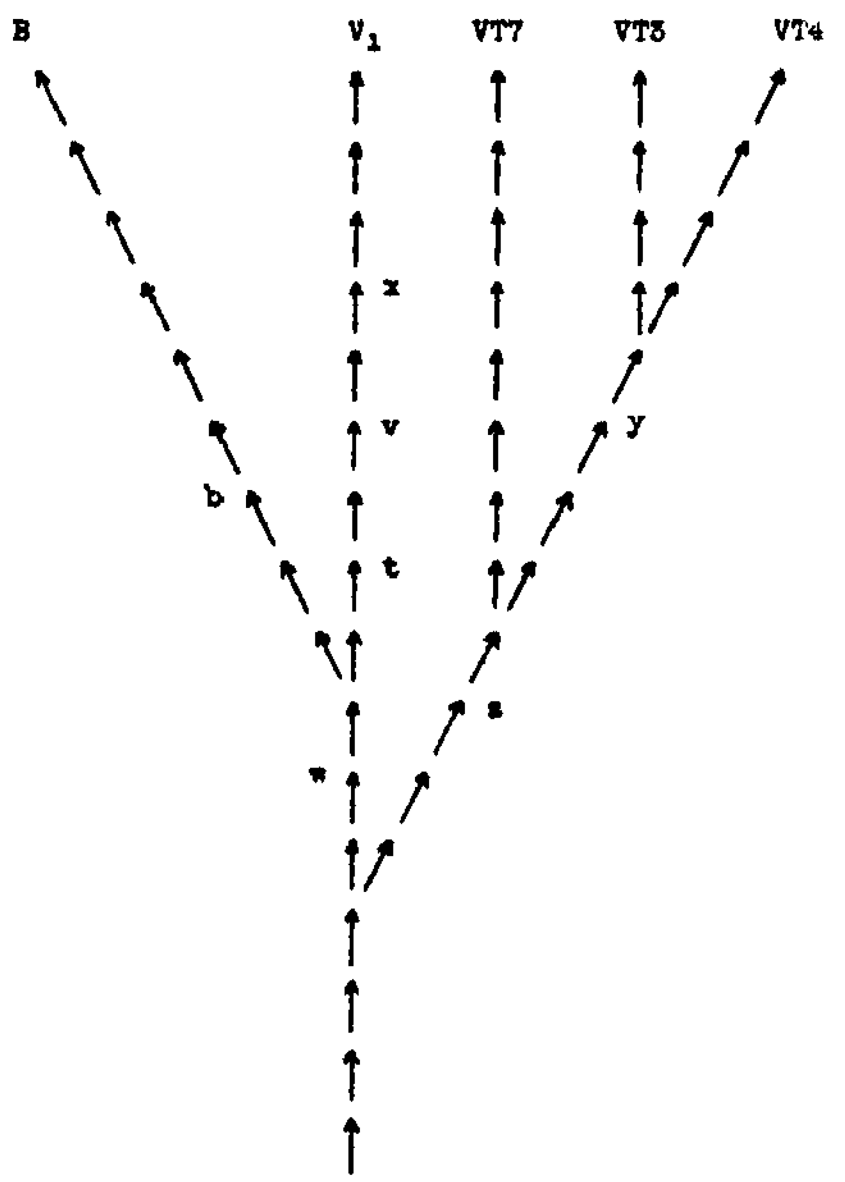


FIGURE 1

Schematic representation of stepwise reactions in the synthesis of essential bacterial metabolites and of elements specific for the adsorption and/or growth of virus.

B = essential bacterial metabolite
V = element(s) specific for virus adsorption and/or growth
b = precursor utilized in synthesis of B only
v = precursor utilized in synthesis of V_1 only
w = mutational block inhibiting formation of precursor common to synthesis of B and V_1
t, x, y, z = mutational blocks inhibiting, respectively, the formation of precursors to the synthesis of V_1; V_1; $VT3$ and $VT4$; $VT7$, $VT3$ and $VT4$.

All strains unable to grow in the minimal medium were resistant to α. This may indicate that the missing factor or factors may be the same for all deficient strains. Not all mutants resistant to α, however, require supplemental factors for growth. This observation suggests that resistance to a virus may be the result of different mutations.

In an attempt to determine what substance or substances the mutants are unable to synthesize, three representative strains were systematically tested in the minimal medium with added vitamins, amino acids, and hydrolyzed casein. None of 8 vitamins or 23 amino acids tested singly or in various combinations, both with and without the addition of hydrolyzed casein, were found to be active under the conditions of any of the tests carried out. So far we have been unable to obtain growth of any of the deficient mutants in a synthetic medium which does not contain nutrient broth or minute amounts of yeast extract.

In a further attempt to determine in what manner the metabolic activities of the mutant strains vary from strain to strain and from those of the original organism a number of qualitative tests of the biochemical characteristics of 14 representative mutants have been carried out. All strains show "typical" *Escherichia coli* reactions and, except for a considerable variation among the cultures in the time required for the full development of the fermentations, etc., in no case have these tests shown any of the strains to have qualitative physiological characteristics which differ from those of the B strain.

Discussion.—The observation that many of the virus-resistant mutants need factors for growth which are not required by the virus-sensitive strain indicates that such a mutant may differ in some profound manner from the parent strain. This is further borne out by the fact that the resistant mutants may exhibit serological properties different from those of the parent organism.[7] This would tend to indicate that one direct effect of the mutation is to either alter the structure of the surface elements to which the virus is specifically adsorbed or to block the synthesis of a specific constituent essential to the reaction of cell and virus.

Each biosynthesis is believed to involve a series of stepwise reactions governed by enzymes which are in turn regulated by the genetic constitution of the organism. The stable heritable mutation may therefore be considered to involve a biochemical change influencing or blocking some step in the synthesis of the altered or missing element or elements necessary for the interaction between cell and virus. The blocked step may be postulated to be one in a series of reactions concerned solely in the synthesis of these specific elements, as at the point x in figure 1. Such a mutation would be one in which the synthesis of a substance non-essential to the growth and reproduction of the organism is blocked. The resulting mutant would not require factors for growth other than those necessary for the parent organism but would be resistant to the virus or viruses for which the specific elements are lacking.

As represented in figure 1, a mutation may occur at reaction w. A block at this point would prevent the synthesis of an intermediate product utilized by the biochemical systems of the cell in the formation of essential metabolic products B and also in the synthesis of specific components V_1 required for the reaction with a virus. In this case the mutant would be resistant and in addition would require factors for growth beyond those needed by the organism from which it arose. This type of mutant could continue to develop, however, if the environment supplied a product which would normally arise at some point *b* in the series of reactions involved in the synthesis of the essential metabolite *B*.

Testing of the mutants for their ability to grow in a minimal medium has indicated that both of these two types may occur, in that some of the resistant mutants require additional factors for growth while others do not.

From a consideration of the hypothesis represented in the diagram the occurrence of a third type of mutant may be anticipated. If virus-resistant mutants were to be isolated in the minimal medium they would of necessity be of a type which requires no factors for growth beyond those needed by the original organism. If a mutation occurred at point *t*, these organisms on transfer to a complete medium may become sensitive to the virus if the medium contains products such as would normally arise at some point *v* and which could then enter into the reactions concerned in the synthesis of the product V_1 which is essential for the specific adsorption and/or growth of the virus.

The diagram may also be useful in interpreting certain findings of Demerec and Fano.[5] These authors, using the same B strain of *Escherichia coli* but employing a larger number of virus strains, have found that mutations may occur which result in resistance not to a single virus but to several virus strains which are totally unrelated on the basis of their morphology and serological reactions.[6,8]

For instance, although the viruses *T3* and *T4* differ serologically and morphologically, Demerec and Fano have found that an organism which is resistant to one is always resistant to the other. In addition, resistance to *T3* and *T4* is facultatively coupled with resistance to *T7* while resistance to *T7* is always coupled with resistance to *T3* and *T4*. Such coupled resistances have been quite generally interpreted as suggesting the existence of relationships among the virus strains. From a consideration of the diagram, however, these types of resistance could arise from mutations at *y* and *z*, respectively. A mutation at *y* would block the synthesis of substances which are necessary for sensitivity to both *T3* and *T4*. A mutational block at *z* would break the chain further back and interfere with the production of substances necessary for sensitivity to all three viruses. The assumption that mutations at points *y* and *z* are the most frequent of those which may affect the sensitivity of the mutants to these three virus strains would allow the hypothesis represented in the diagram to explain the findings of Demerec and Fano.

The linkage of the resistances to *T3* and to *T4* may therefore be due not to a similarity between these two viruses but rather to the occurrence of a common precursor in the biochemical reactions of the organism which lead to the synthesis of the substances essential for the adsorption and/or growth of these two viruses.

Summary.—Virus-resistant mutants have been isolated from a virus-sensitive strain of *Escherichia coli* utilizing the selective action of three strains of bacterial viruses active on the parent organism. The mutants differ from the original strain by their resistance to either one, any two, or all three viruses, in all possible combinations. Morphological studies and qualitative physiological tests show all strains tested to have characteristics similar to the original organism. Of the 57 mutants isolated in pure culture, 28 were unable to grow in a basal medium but were able to grow when 0.005 per cent yeast extract was added to this medium. Three representatives were systematically tested with various combinations of 8 vitamins, 23 amino acids, and hydrolyzed casein. None of these substances, added to the basal medium singly or in combination, would support growth of the strains tested.

[*] Fellow in the Medical Sciences of the NATIONAL RESEARCH COUNCIL.

The author wishes to thank Dr. M. Delbrück for his unfailing counsel during these investigations and the preparation of this manuscript. Grateful acknowledgment is also made to Miss A. B. Martin for assistance in carrying out the morphological and qualitative physiological tests.

[1] Bail, O., *Z. Immun. Forsch.*, **38,** 57–164 (1923); D'Herelle, F., *The Bacteriophage and Its Behavior*, Baltimore, Williams & Wilkins, 1926; Gratia, A., *Jour. Exp. Med.*, **34,** 115–131 (1921); Burnet, F. M., *Jour. Path. Bact.*, **32,** 15–42 (1929).

[2] Luria, S. E., and Delbrück, M., *Genetics*, **28,** 491–511 (1943).

[3] Delbrück, M., and Luria, S. E., *Arch. Biochem.*, **1,** 111–141 (1942).

[4] Luria, S. E., Delbrück, M., and Anderson, T. F., *Jour. Bact.*, **46,** 57–77 (1943).

[5] Demerec, M., and Fano, U., *Genetics*, "Bacteriophage-Resistant Mutants in *Escherichia coli*," in press.

[6] Beadle, G. W., and Tatum, E. L., these PROCEEDINGS, **27,** 499–506 (1941); Tatum, E. L., and Beadle, G. W., *Growth, Fourth Growth Symposium*, Vol. 6, 27–35 (1942).

[7] Burnet, F. M., *Jour. Path. Bact.*, **33,** 647–664 (1930); Delbrück, M., unpublished data.

[8] Anderson, T. F., Delbrück, M., and Demerec, M., "Types of Morphology Found in Bacterial Viruses," paper presented before the Electron Microscope Society of America, Chicago (Nov. 15–17, 1944).

X-RAY INDUCED GROWTH FACTOR REQUIREMENTS IN BACTERIA*

By C. H. Gray and E. L. Tatum

School of Biological Sciences, Stanford University

Communicated October 20, 1944

The specific growth factor requirements of bacteria result from the inability of these organisms to synthesize particular essential substances from simpler constituents.[1, 2] In some instances such deficiencies have been shown to result from an inability to carry out definite chemical reactions in a biosynthesis.[3] These defects in synthesis are characteristic of each cell of a given strain of bacteria, and are in addition transmitted to each descendant cell. Similar defects in certain biosyntheses have been induced in the ascomycete *Neurospora* by irradiation and have been shown to result from mutations in specific genes.[4, 5] This suggests that if biochemical reactions in bacteria are likewise under genetic control, mutations in certain "genes" in these organisms should lead to corresponding growth factor deficiencies in the modified strains. Treatment with x-rays might be expected to induce such "gene" mutations in bacteria. This paper presents evidence that x-ray treatment of bacteria does produce mutant types, presumably genic, with altered growth factor requirements.[6]

Two species of bacteria have been used in this investigation: *Escherichia coli* (K-12) and *Acetobacter melanogenum* (M.A. 6.1). The minimal medium used for *E. coli* contained per liter: NH_4CL, 5.0 g.; NH_4NO_3, 1.0 g.; Na_2SO_4, 2.0 g.; $MgSO_4 \cdot 7H_2O$, 0.1 g.; $CaCl_2$, trace; K_2HPO_4, 3.0 g.; KH_2PO_4, 1.0 g.; 1 cc. trace element solution[7]; glucose, 5.0 g.; and vitamin-free *l*-asparagine, 1.5 g. (optional). *A. melanogenum* has been found to grow well on a minimal medium adjusted to pH 6 with acetic acid and containing per liter: K_2HPO_4, 0.5 g.; KH_2PO_4, 0.5 g.; $MgSO_4 \cdot 7H_2O$, 0.2 g.; $FeSO_4$, $MnSO_4$ and NaCl, 0.01 g. each; $(NH_4)_2SO_4$, 10 g.; glycerol, 50 g.; thiamin chloride hydrochloride, 1 mg.; pantoyl *l*-lactone, 1 mg.; nicotinic acid, 1 mg.; *p*-aminobenzoic acid, 0.1 mg.; and *l*-proline, 0.2 g. (optional).

The bacteria were treated with x-rays produced in a Westinghouse "Duocondex" unit containing an industrial type oil-cooled tube operated at 160 kv. and 25 ma. Twenty-four or 48-hour-old liquid cultures of bacteria were exposed to a total irradiation of 180,000 r units given at a rate of 4000 per minute. Preliminary tests had shown that at this dosage over 99.99 per cent of the cells were killed. The irradiated suspensions were then incubated for four hours in yeast extract broth on a shaker. This procedure was adopted to increase the probability of any individual cells altered by the treatment giving rise to homogeneous colonies. Samples of these suspensions were plated out in complete agar medium containing

0.3 per cent yeast extract, 0.5 per cent peptone and 0.5 per cent glucose. Dilutions were made so that approximately 100 colonies were obtained per plate. The plates were incubated for several days and single well-isolated colonies were then picked and transferred to slants of complete medium. These cultures were tested for their ability to grow in liquid minimal media. All strains which grew in this first test were transferred to a second tube of minimal medium in order to decrease the possibility of any mechanical carry-over of essential growth factors. Strains of these bacteria which failed to grow on the minimal media, but which grew well on complete medium, were carried through one or two more successive platings and reisolated. They were then tested for their growth factor requirements in a manner analogous to that previously described for the characterization of *Neurospora* mutant strains.[4]

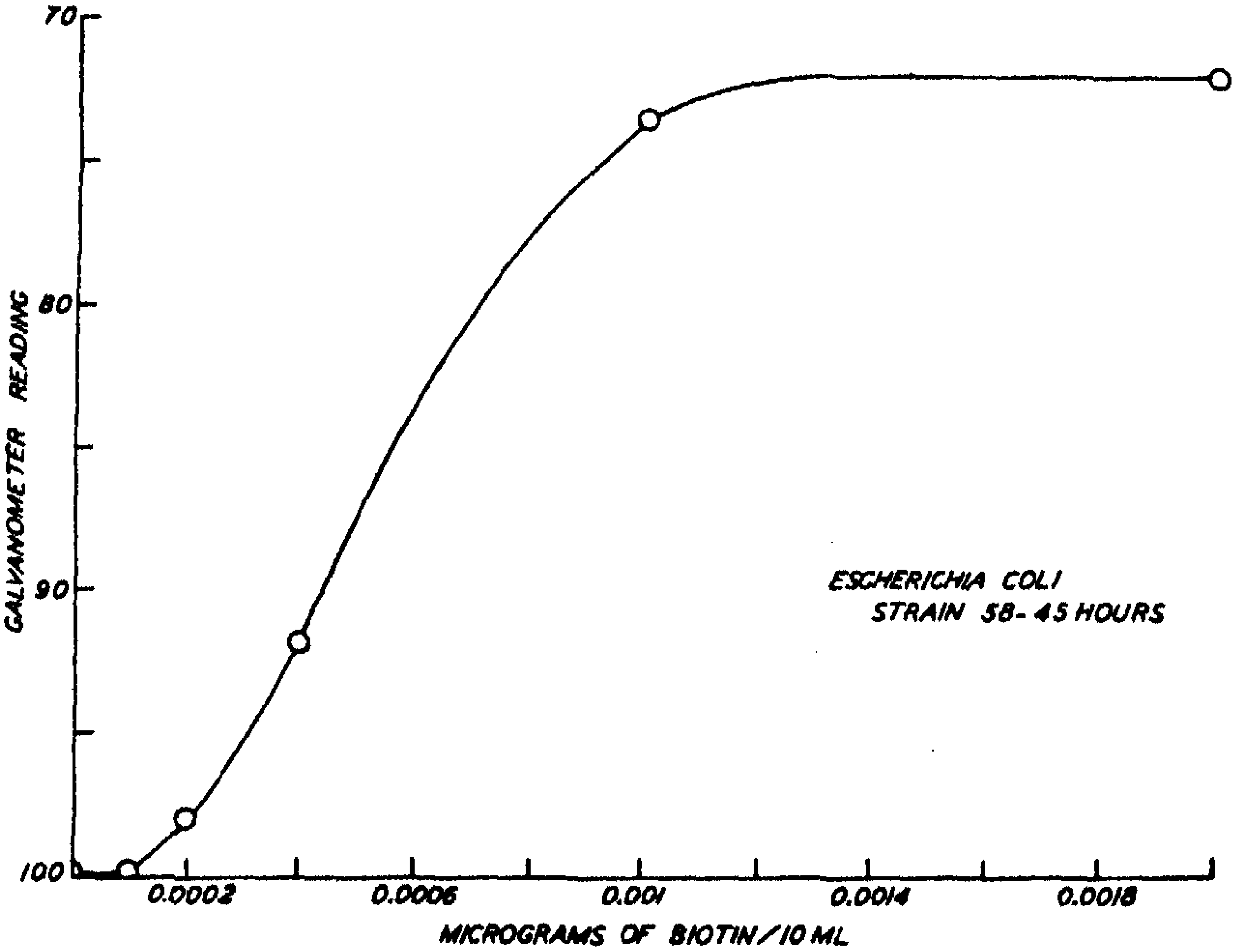


FIGURE 1

Growth of *E. coli*, strain 58, cultured at 25°C. on minimal medium with various concentrations of biotin. Under the same conditions *E. coli* (K12) gave a reading of 64 on minimal medium.[3]

From 800 isolations two strains of *Escherichia coli* have been obtained which are distinguished from the original strain by their inability to grow in the minimal *E. coli* medium. One of these, strain 58, requires the addition of biotin, and the other, strain 679, requires the amino acid threonine. Not only does each of these derived strains respond to the addition of the specific substance required, but it responds only to that substance. In each case the growth response is a function of the amount of growth factor supplied (Figs. 1 and 2).

From 575 isolations four strains of *Acetobacter melanogenum* have been obtained which fail to grow in the appropriate minimal medium. Three of these grow normally when supplied with the required substances. Strain 90 responds to the addition of either glycine or *l*-serine (Fig. 3). Strain 92 requires adenine or adenosine (Fig. 4). Strain 236 requires glycine and will not respond to the addition of serine (Fig. 5). The requirement of strain 318 has not been completely analyzed as yet. It responds to the addition of leucine, but does not grow as well as the original strain with any concentration of leucine so far tested.

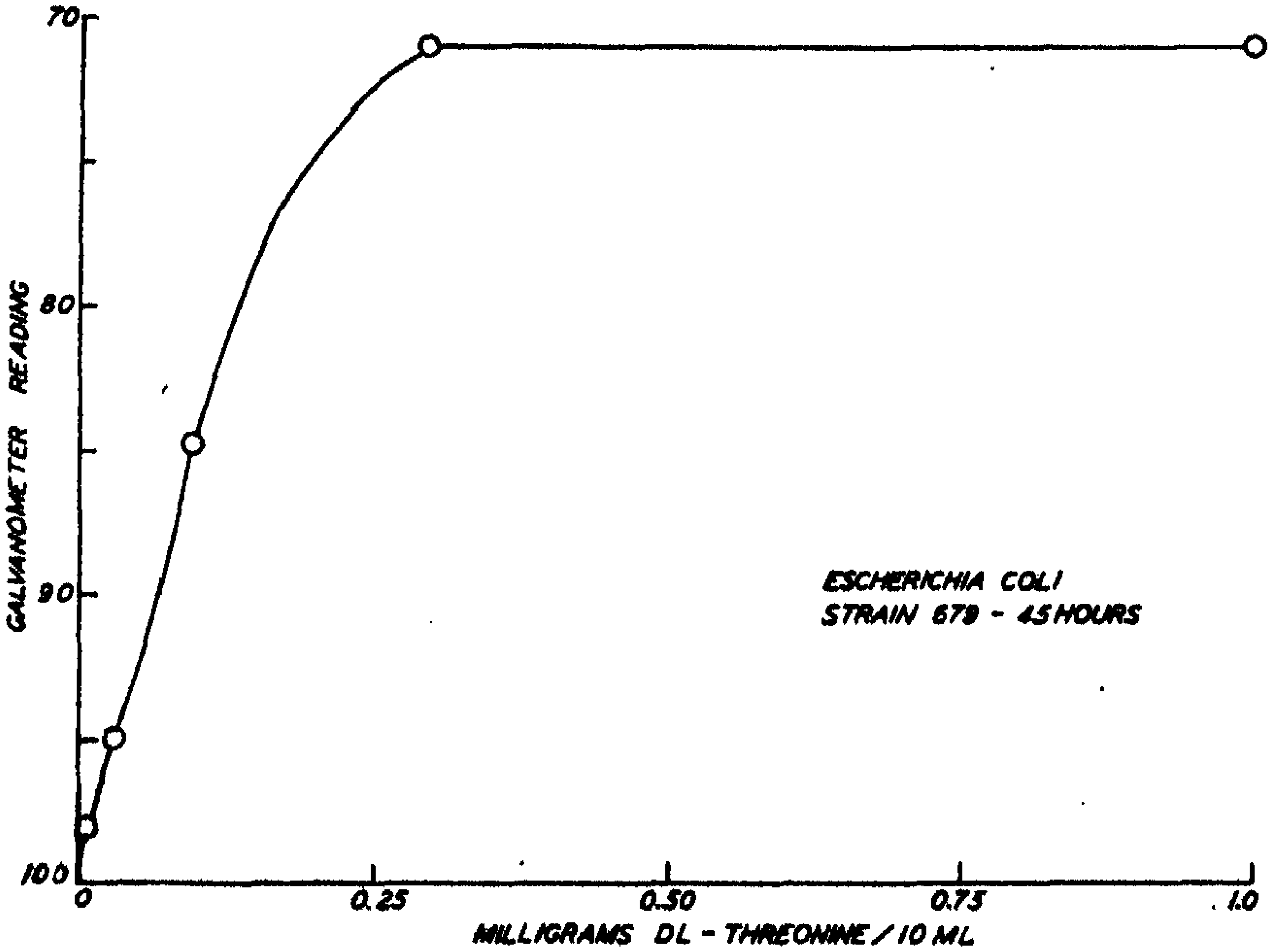


FIGURE 2

Growth of *E. coli*, strain 679, cultured at 25°C. on minimal medium with various concentrations of threonine.[5]

The general characteristics of each culture have been compared with those of the original strains. All cultures were indistinguishable from the parent strains both macroscopically and microscopically. The derived strains of *E. coli* were gram-negative, gave positive indole and methyl red tests, and produced gas. The variant strains of *A. melanogenum* resembled *A. melanogenum* M.A. 6.1 in the production of acid, of pigment, and of di-hydroxyacetone from glycerol, and in being gram-negative. Each of these derived strains had the same requirements as the original *A. melanogenum* for thiamin, *p*-aminobenzoic acid, nicotinic acid and pantoyl-*l*-lactone, and each was able to use either ammonia or *l*-proline as a nitrogen source.

All of the available evidence is consistent with the view that the new strains of *E. coli* and *A. melanogenum* were produced as a result of the x-ray treatment. No variant strains of either organism were obtained following treatment with ultraviolet light, although as many isolations were made and tested as after x-ray treatment. It therefore seems improbable that the variant strains obtained after x-raying could have been present in the original stock or that they could have resulted from the selection of spontaneous variants. It is unlikely that either of these species would be present among the airborne organisms in the laboratory. Even more improbable would be the presence of a number of strains of these species which differed from the original strains under investigation only in having different specific growth factor requirements. These considerations

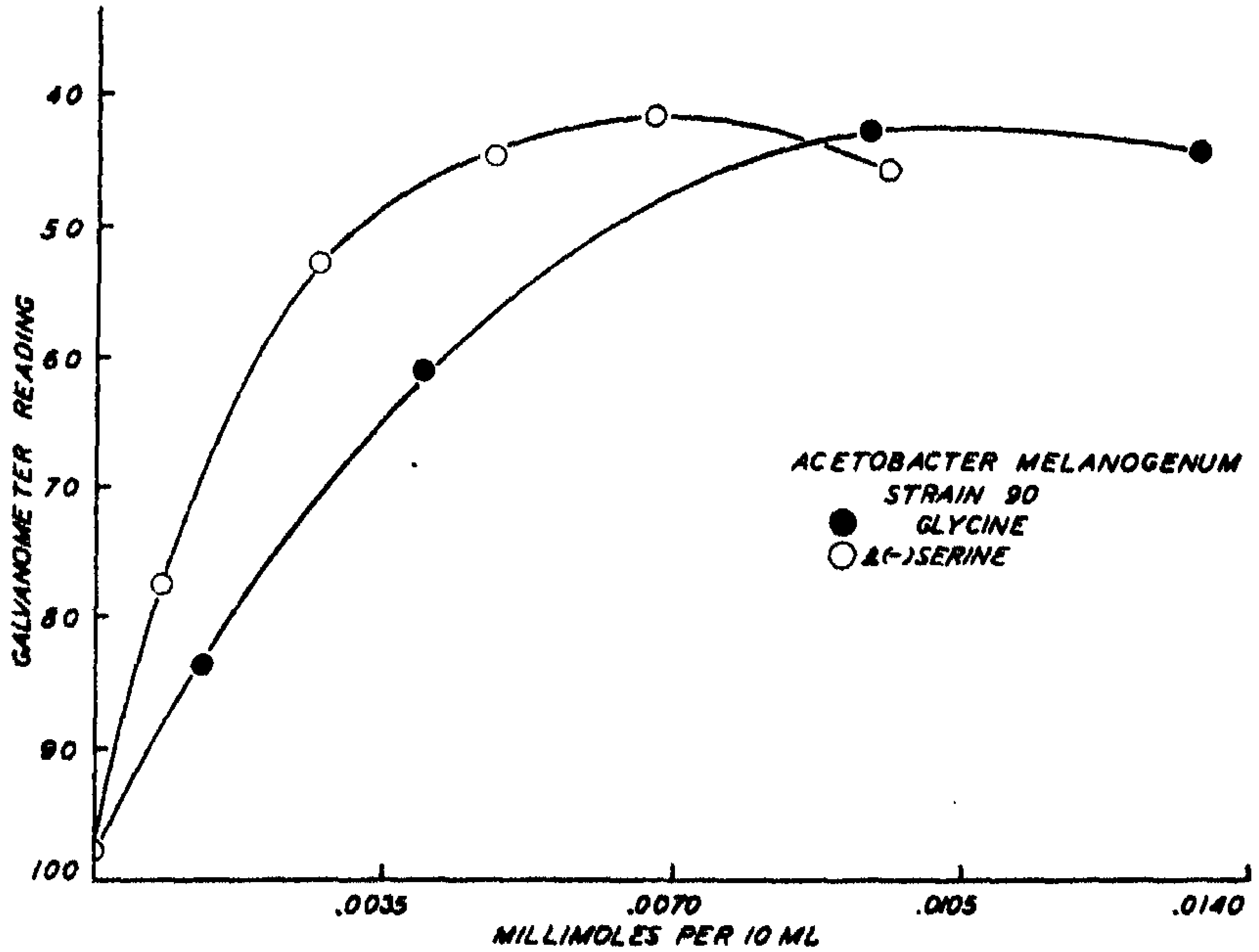


FIGURE 3

Growth of *A. melanogenum*, strain 90, after 72 hours of constant shaking at 25°C. in minimal medium with various concentrations of either glycine or serine. Under the same conditions *A. melanogenum* (M.A. 6.1) gave a reading of 44 on minimal medium.[8]

strongly suggest that the x-ray treatment was responsible for the production of the mutant strains obtained in *E. coli* and *A. melanogenum*.

Definite proof of the existence of genes in bacteria cannot be obtained by the methods of classical genetics since no sexual phases in their life cycles have yet been demonstrated. For the same reason it is at present impossible to distinguish between mutations in nuclear genes and in cytoplasmic determiners or plasmagenes[9] in bacteria. The term "gene" can therefore be used in connection with bacteria only in a general sense. Nevertheless, the rôle of genes in bacterial variation has been suggested and the existence in bacteria of discrete nuclear material, perhaps analogous to chromosomes, has been demonstrated in certain forms.[10] The production by x-ray treatment of the new strains of two species of bacteria described in this paper offers indirect support for the existence of genes in bacteria. X-rays are known to increase the rate of gene mutations in other organisms and the mutation or inactivation of single genes seems the most reasonable explanation for the origin of the variant strains obtained since each differs from the original strain in having lost the ability to synthesize specific growth factors. These deficiencies are similar to those which have been

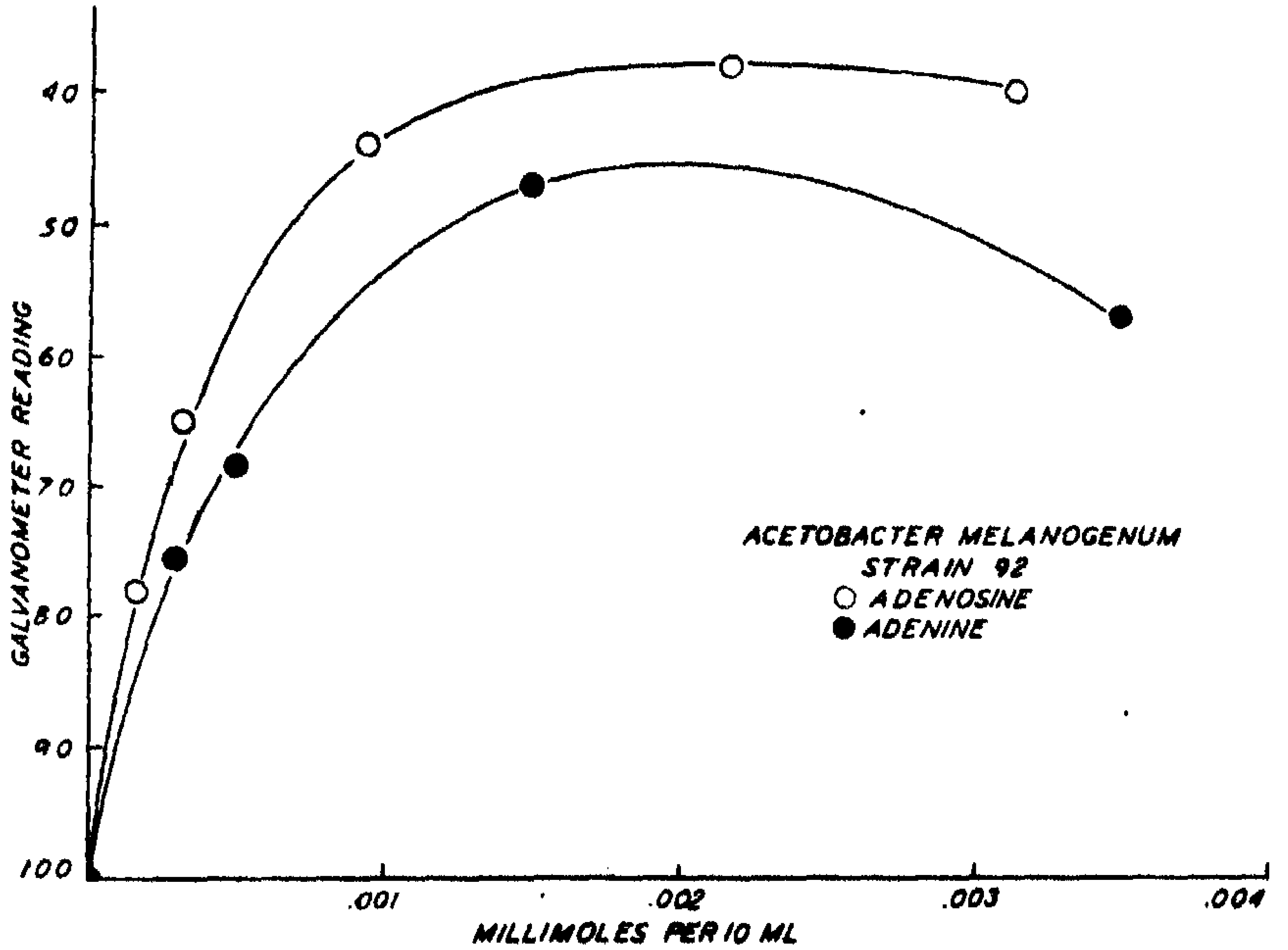


FIGURE 4

Growth of *A. melanogenum*, strain 92, after 72 hours of constant shaking at 25°C. in minimal medium with various concentrations of either adenine or adenosine.[8]

shown to be associated with mutations in single nuclear genes in *Neurospora.*[4, 5]

Summary.—Strains of two species of bacteria have been obtained following x-ray treatment which are characterized by their inability to carry out specific biochemical reactions. Two strains of *Escherichia coli* differ from the parent strain, one by requiring biotin, and the other by requiring threonine. Four mutant strains of *Acetobacter melanogenum* have been obtained. The first requires a supplement of serine or glycine, the second requires adenine or adenosine, the third glycine, and the last leucine.

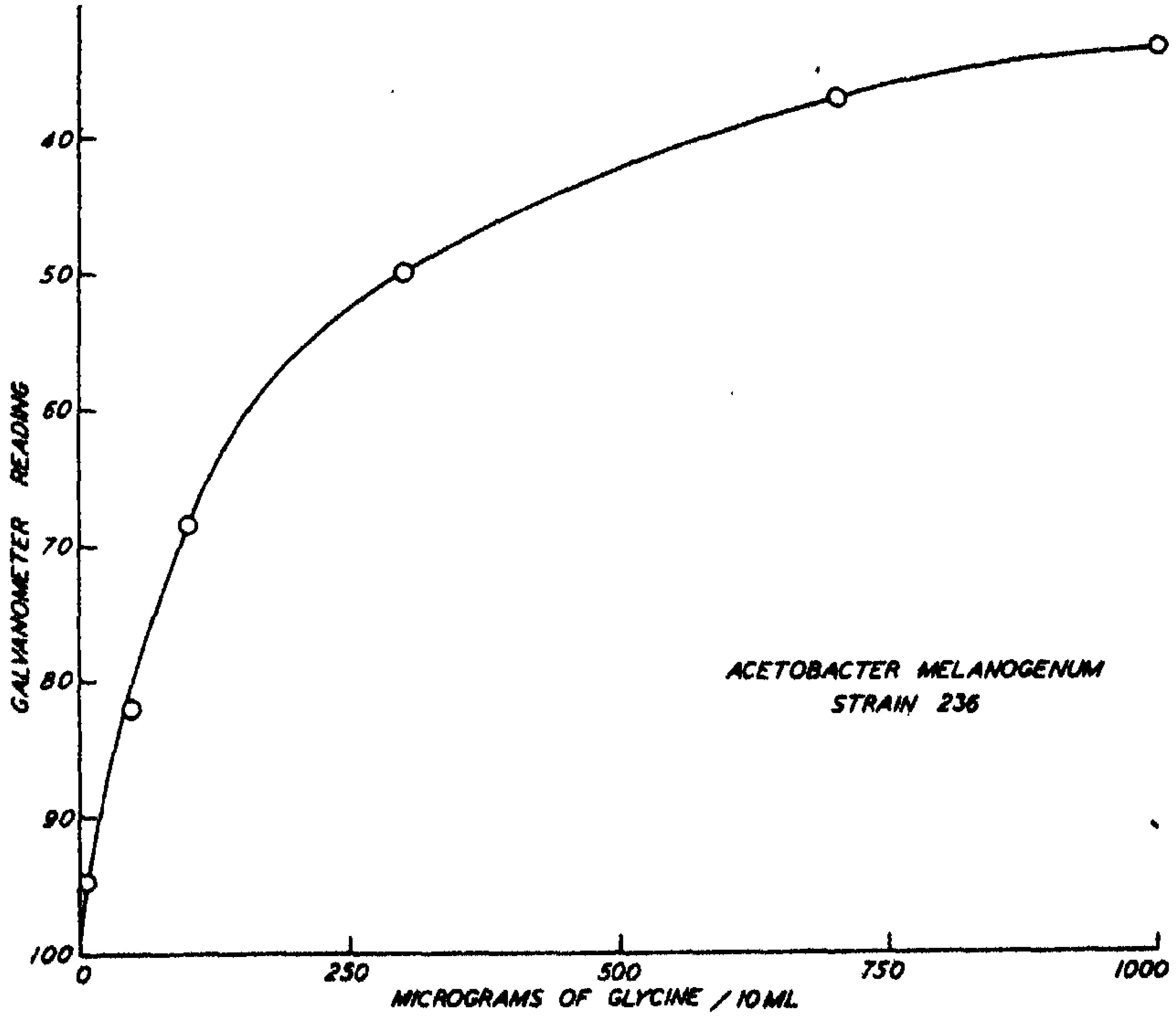


FIGURE 5

Growth of *A. melanogenum*, strain 236, after 72 hours of constant shaking at 25°C. in minimal medium with various concentrations of glycine.[8]

With each strain there is little or no growth in the absence of the required substance as contrasted to the excellent growth of the parent strain in the appropriate minimal medium. The growth of each strain is a function of the concentration of the required substance.

These results prove that the capacity of bacteria for carrying out specific biochemical reactions can be modified by x-ray treatment, and suggest that biosyntheses in bacteria are controlled by specific genes.

[*] Work supported in part by a grant from the Rockefeller Foundation. Preliminary report read at the 39th Annual Meeting of the Botanical Society of America at Cleveland, Ohio, September 12, 1944. The authors are indebted to Merck & Co., Inc., Rahway, N. J., for samples of pantoyl-*l*-lactone and biotin; to Dr. C. E. Clifton, Dept. of Bacteriology, Stanford University, for the strain of *E. coli;* and to Dr. C. B. van Niel, Hopkins Marine Station, Stanford University, for the strain of *A. melanogenum.*

[1] Knight, B. C. J. G., "Bacterial Nutrition," *Med. Res. Coun. (Brit.) Spec. Rep. Ser.*, 210 (1936).

[2] Lwoff, A., *Ann. Inst. Pasteur*, **61**, 580 (1939).

[3] Schlenk, F., and Gingrich, W., *Jour. Biol. Chem.*, **143**, 295 (1942).

[4] Beadle, G. W., and Tatum, E. L., these PROCEEDINGS, **27**, 499 (1941).

[5] Tatum, E. L., *Annual Rev. of Biochem.*, Stanford University, Vol. 13, 667 (1644).

[6] A number of similar mutant strains of *E. coli* have recently been described by Lampen, J. O., Jones, M. J., and Roepke, R. R., before the Division of Biological Chemistry at the New York Meeting of the American Chemical Society, September 14, 1944.

[7] As used for *Neurospora*, see Horowitz, N. H., and Beadle, G. W., *Jour. Biol. Chem.*, **150**, 325 (1943).

[8] Growth determined turbidimetrically with an Evelyn Photoelectric Colorimeter, using 10 ml. culture with a 540 μ filter. Galvanometer reading of 100 represents no growth, whereas a reading of 0.0 represents a completely opaque solution.

[9] See Darlington, C. D., *Nature*, **154**, 164 (1944).

[10] For reviews of this topic see Lewis, I. M., *Bact. Reviews*, **5**, 181 (1941), and Knaysi, G. *Elements of Bacterial Cytology*, Comstock Publishing Co., 1944, 209 pp.

INDEX

Abnormality, specific (Dunn and Gluecksohn-Schoenheimer), 173.
Activity (Wald and Jackson), 255.
Adaptive enzymes (Lindegren, Spiegelman and Lindegren), 346.
Alleles, oppositional (Atwood), 69.
ALBERT, A. A. The Minimum Rank of a Correlation Matrix, 144.
——. The Matrices of Factor Analysis, 90.
Analysis, rational diophantine (Bell), 355.
ANDERSON, E. H. Incidence of Metabolic Changes among Virus-Resistant Mutants of a Bacterial Strain, 396.
Antibiotic activity of lichens (Burkholder, Evans, McVeigh and Thornton), 250.
Antibiotic substances (Waksman and Reilly), 99.
Antibodies (Emerson), 179.
Apomixis in guayule (Esau), 352.
Arithmetic in a group ring (Brauer), 109.
Associations, chromosome (Bergner), 302.
ATWOOD, S. S. The Behavior of Oppositional Alleles in Polyploids of *Trifolium repens*, 69.
Automorphism groups (Chevalley), 274.

Bacteria (Gray and Tatum), 403.
Bacterial strains (Anderson), 396.
BALL, E., AND BOELL, E. J. Respiratory Rates of the Shoot Tips and Maturing Tissues in *Lupinus albus* and *Tropaeolum majus*, 45.
BALLENTINE, R. See McElroy, W. D., 377.
BARAJAS, A. Birkhoff's Theory of Gravitation and Einstein's Theory for Weak Fields, 54.
BARKER, H. A. On the Rôle of Carbon Dioxide in the Metabolism of *Clostridium thermoaceticum*, 88.
BATEMAN, H. Note on the Function $F(a, b; c - n; z)$, 28.
BELL, E. T. A Method in Rational Diophantine Analysis, 355.
BERGNER, A. D. Chromosome Associations in Tetraploid Hybrids between Prime Type 1 and Prime Type 2 in *Datura stramonium*, 302.
Biological effects of X-rays (Bless), 118.
Bioluminescence, mechanics of (McElroy and Ballentine), 377.
Biosynthesis of tryptophane (Tatum and Bonner), 30.
BIRKHOFF, G. D. Flat Space-Time and Gravitation, 324.
Birkhoff's theory of gravitation (Barajas), 54; (Weyl), 205.
BLESS, A. A. The Biological Effects of X-Rays as a Function of Intensity, 118.
BLOCH, R. See Sinnott, E. W., 387.
Blood types (Sawin, Griffith and Stuart), 217.
BOELL, E. J. See Ball, E., 45.
BONNER, D. See Tatum, E. L., 30.
Brachy (T/T) (Gluecksohn - Schoenheimer), 134.
BRADLEY, M. See Marshak, A., 231.
BRAUER, R. On the Arithmetic in a Group Ring, 109.
Bundles, Sphere (Chern), 269.
BURKHOLDER, P. R., EVANS, A. W., McVEIGH, I., AND THORNTON, H. K. Antibiotic Activity of Lichens, 250.

Carbon dioxide (Barker), 88.
Cartography, general (Kasner and DeCicco), 211.
CASTLE, W. E. Linkage of Waltzing in the Rat, 226.
CASTLE, W. E., AND KING, H. D. Linkage Studies of the Rat (*Rattus Norvegicus*) IV, 79.
Characterization, group-theoretic (Mendelsohn), 279.
Characters of groups and semi-groups (Hille), 58.
CHERN, S. Laplace Transforms of a Class of Higher Dimensional Varieties in a Projective Space of *n* Dimensions, 95.
——. Integral Formulas for the Characteristic Classes of Sphere Bundles, 269.
CHEVALLEY, C. On Groups of Automorphisms of Lie Groups, 274.
Chiasmata, reciprocal (Cooper), 50.
Chick (Gluecksohn-Schoenheimer), 134.
Chromosome associations (Bergner), 302.
Chromosome number (Marshak and Bradley), 231.
Chromosome, sex (Cooper), 50.
Chromosomes (Sparrow), 147; (MacKnight & Cooper), 384.
Chrysogenum group (Waksman and Reilly), 99.
Clostridium thermoaceticum (Barker), 88.
Clusters, globular (Shapley), 61.
Coelom of the chick (Gluecksohn-Schoenheimer), 134.
Collineation, projective (Mendelsohn), 279.
Compressibility of media (Murnaghan), 244.
Conformal maps (Kasner and DeCicco), 162.
Congruences, trinomial (Vandiver), 368.
COOPER, K. W. Invalidation of the Cytological Evidence for Reciprocal Chiasmata in the Sex Chromosome Bivalent of Male Drosophila, 50.

Cooper, K. W. See MacKnight, R. H., 384.
Correlation matrix (Albert), 144; (Thurstone), 129.
Curve, Soper's epidemic (Wilson and Worcester), 37.
Curves, scale (Kasner and DeCicco), 162, 211.
Cylinder, elastic (Murnaghan), 382.
Cytoplasmic pattern (Sinnott and Bloch), 387.

Datura stramonium (Bergner), 302.
DeCicco, J. See Kasner, E., 162, 211.
Degenerate form of Einstein's theory (Weyl), 205.
Deprivation, nutritional (Wald and Jackson), 255.
Development of the eye (Steinberg), 5.
Differentiation of xylem strands (Sinnott and Bloch), 387.
Diophantine analysis (Bell), 355.
Displacements, elastic (Thomas), 140.
Distances of globular clusters (Shapley), 61.
Distortion tensor (Thomas), 140.
Dobzhansky, T., and Mayr, E. Experiments on Sexual Isolation in Drosophila. I. Geographic Strains of *Drosophila willistoni*, 238.
Dobzhansky, T., and Streisinger, G. Experiments on Sexurl Isolation in Drosophila. II. Geographic Strains of *Drosophila prosaltans*, 340.
——. Experiments on Sexual Isolation in Drosophila. III. Geographic Strains of *Drosophila sturtevanti*, 335.
Dominant mutation in Maize (Stadler), 123.
Drosophila (Dobzhansky and Streisinger), 340; (Dobzhansky), 335.
Drosophila, male (Cooper), 50.
Drosophila melanogaster (Steinberg), 5.
Drosophila miranda (MacKnight and Cooper), 384.
Drosophila prosaltans (Dobzhansky and Streisinger), 340.
Drosophila, sexual isolation in (Dobzhansky and Mayr), 238.
Drosophila willistoni (Dobzhansky and Mayr), 238.
Dunn, L. C., and Gluecksohn-Schoenheimer, S. A Specific Abnormality Associated with a Variety of Genotypes, 173.
Dynamical systems (Ingram), 370.

Einstein's theory of gravitation (Barajas), 54; (Weyl), 205.
Elastic cylinder (Murnaghan), 382.
Elastic displacements (Thomas), 140.
Embryos, mouse (Gluecksohn-Schoenheimer), 134.
Emerson, S. The Induction of Mutations by Antibodies, 179.
Enzymatic adaptation (Spiegelman, Lindegren and Hedgecock), 13.
Enzymes, adaptive (Lindegren, Spiegelman and Lindegren), 346.
Ephrussi, B., and Sutton, E. A Reconsideration of the Mechanism of Position Effect, 183.
Epidemic curve (Wilson and Worcester), 37, 264.
Equations, integral (Ingram), 370.
Equilibrium in genic materials (Jones), 82.
Errata (G. W. Mackey) and (H. S. Uhler), 24; (Bateman), 44
Esau, K. Apomixis in Guayule, 352.
Evans, A. W. See Burkholder, P. R., 250.
Extragalactic studies (Shapley), 61.
Eye, development of (Steinberg), 5.
Eyeless mutants (Steinberg), 5.

Factor analysis, matrices of (Albert), 90.
Factoring correlation matrix (Thurstone), 129.
Falco, E. A. See Hitchings, G. H., 294.
Fermat's last theorem (Vandiver), 362, 368.
Ferns, common (King), 155.
Field theory (Vandiver), 362.
Fields, weak (Barajas), 54.
Finite field theory (Vandiver), 362.
Fishes, production of guanine and vision in (Sumner), 285.
Flat space-time (Birkhoff), 324.
Flows, parallel (Lin), 316.
Function $F(a, b; c - n; z)$ (Bateman), 28.

Galactic studies (Shapley), 61.
Gasteria, pericentric inversion in (Giles), 1.
General ransform (Reed), 169.
Genes, normal and mutant (Stern and Heidenthal), 197.
Genetic linkage of blood types (Sawin, Griffith and Stuart), 217.
Genetically controlled yeast populations (Spiegelman, Lindegren and Hedgecock), 13.
Genic materials (Jones), 82.
Genotypes, variety of (Dunn and Gluecksohn-Schoenheimer), 173.
Geographic strains of *Drosophila prosaltans* and *sturtevanti* (Dobzhansky and Streisinger), 340; (Dobzhansky), 335.
Geographic strains of *Drosophila willistoni* (Dobzhansky and Mayr), 238.
Giles, N. H., Jr. A Pericentric Inversion in Gasteria Resulting in Apparent Iso-Chromosomes at Meiosis, 1
Girella nigricans extract (Hitchings and Falco), 294.
Globular clusters (Shapley), 61.

Gluecksohn-Schoenheimer, S. The Development of Normal and Homozygous Brachy (T/T) Mouse Embryos in the Extraembryonic Coelom of the Chick, 134.
Gluecksohn-Schoenheimer, S. See Dunn, L. C., 173.
Goldschmidt, R. On Spontaneous Mutation, 297.
Goldschmidt, R., and Hannah, A. One-Band Inversion, 299.
Gravitation (Barajas), 54; (Weyl), 205; (Birkhoff), 324.
Gray, C. H., and Tatum, E. L. X-Ray Induced Growth Factor Requirements in Bacteria, 403.
Griffith, M. A. See Sawin, P. B., 217.
Group ring (Brauer), 109.
Groups (Miller), 25; (Hille), 58; (Miller), 114; (Chevalley), 274; (Miller), 275; (Mendelsohn), 279; (Miller), 359.
Group-theoretic characterizations (Mendelsohn), 279.
Growth-delaying effect of ultraviolet radiation (Luria), 392.
Growth factors (Gray and Tatum), 403.
Guanase, specificity of (Hitchings and Falco), 294.
Guanine (Sumner), 285; (Hitchings and Falco), 294
Guayule, apomixis in (Esau), 352.

Hannah, A. See Goldschmidt, R., 299.
Haro, G. Note on the Discovery of Red Stars, 247.
Hedgecock, L. See Spiegelman, S., 13.
Heidenthal, G. See Stern, C., 197.
Hille, E. On the Theory of Characters of Groups and Semi-Groups in Normed Vector Rings, 58.
Hitchings, G. H., and Falco, E. A. The Identification of Guanine in Extracts of *Girella nigricans*. The Specificity of Guanase, 294.
Hybrids, tetraploid (Bergner), 302.

Indole (Tatum and Bonner), 30.
Induction of mutations by antibodies (Emerson), 179.
Ingram, W. H. On the Integral Equations of Continuous Dynamical Systems, 370.
Inheritance, Mendelian (Lindegren, Spiegelman and Lindegren), 346.
Inhibition, X-ray (Marshak and Bradley), 231.
Integral equations (Ingram), 370.
Integral formulas (Chern), 269.
Intensity of X-rays (Bless), 118.
Inversion, one-band (Goldschmidt and Hannah), 297.
Inversion, pericentric (Giles), 1.
Iso-chromosomes (Giles), 1.
Isolation, sexual (Dobzhansky and Streisinger), 340; (Dobzhansky), 335.

Jackson, B. See Wald, G., 255.
Jones, D. F. Equilibrium in Genic Materials, 82.

Kasner, E., and DeCicco, J. Scale Curves in Conformal Maps, 162.
——. Scale Curves in General Cartography, 211.
King, A. L. The Spore Discharge Mechanism of Common Ferns, 155.
King, H. D. See Castle, W. E., 79.

Laplace transforms (Chern), 95.
Lebesgue space (Segal), 165.
Lichens, antibiotic activity of (Burkholder, Evans, McVeigh and Thornton), 250.
Lie groups (Chevalley), 274.
Lin, C. C. On the Stability of Two-Dimensional Parallel Flows, 316.
Lindegren, C. C. See Spiegelman, S., 13.
Lindegren, C. C., Spiegelman, S., and Lindegren, G. Mendelian Inheritance of Adaptive Enzymes, 346.
Lindegren, G. See Lindegren, C. C., 346.
Linkage, genetic (Sawin, Griffith and Stuart), 217.
Linkage of waltzing in rats (Whittinghill), 221; (Castle), 226.
Linkage studies (Castle and King), 79.
Lobe^3, lobe^4, lobe^5 (Steinberg), 5.
Lupinus albus (Ball and Boell), 45.
Luria, E. E. A Growth-Delaying Effect of Ultraviolet Radiation, 392.

Mackey, G. W., errata, 24.
McElroy, W. D., and Ballentine, R. The Mechanics of Bioluminescence, 377.
MacKnight, R. H., and Cooper, K. W. The Synapsis of the Sex Chromosomes of *Drosophila miranda* in Relation to Their Directed Segregation, 384.
McVeigh, I. See Burkholder, P. R., 250.
Maize, dominant mutation in (Stadler), 123.
Male Drosophila (Cooper), 50.
Maps, conformal (Kasner and DeCicco), 162.
Marshak, A., and Bradley, M. X-Ray Inhibition of Mitosis in Relation to Chromosome Number, 231.
Matrix, correlation (Thurstone), 129; (Albert), 144.
Maturing tissues (Ball and Boell), 45.
Mayr, E. See Dobzhansky, T., 238.
Mechanical self-excited oscillations (Minorsky), 308.

Mechanics of bioluminescence (McElroy and Ballentine), 377.
Mechanism of enzymatic adaptation (Spiegelman, Lindegren and Hedgecock), 13.
Mechanism of position effect (Ephrussi and Sutton), 183.
Media (Murnaghan), 244.
Meiosis (Giles), 1.
Meiotic chromosomes (Sparrow), 147.
Mendelian inheritance (Lindegren, Spiegelman and Lindegren), 346.
MENDELSOHN, N. S. A Group-Theoretic Characterization of the General Projective Collineation group, 379.
Mersenne number M_{157} (Uhler), 314.
Metabolic changes (Anderson), 396.
Metabolism of *Clostridium thermoaceticum* (Barker), 88.
Matrices of factor analysis (Albert), 90.
MILLER, G. A. Relative Numbers of Non-Invariant Operators in a Group, 25.
——. Possible Numbers of Non-Invariant Operators of a Group, 114.
——. Groups Containing Less than Twenty-Eight Non-Invariant Operators, 275.
——. Groups Involving a Small Number of Sets of Conjugate Operators, 359.
Minimum rank of correlation matrix (Albert), 144.
MINORSKY, N. On Mechanical Self-Excited Oscillations, 308.
Mitosis (Marshak and Bradley), 231.
Mouse embryos (Gluecksohn-Schoenheimer), 134.
MURNAGHAN, F. D. The Compressibility of Media under Extreme Pressures, 244.
——. On the Theory of the Tension of an Elastic Cylinder, 382.
Mutant genes (Stern and Heidenthal), 197.
Mutants, eyeless (Steinberg), 5.
Mutants, virus-resistant (Anderson), 396.
Mutation, spontaneous (Goldschmidt), 297.
Mutations (Emerson), 179.
Mutations, specific (Sturtevant), 176.

Non-invariant operators (Miller), 25, 114, 275.
Normal genes (Stern and Heidenthal), 197.
Normed vector rings (Hille), 58.
Nucleic acid cycle (Sparrow), 147.
Number, Mersenne M_{157} (Uhler), 314.
Nutritional deprivation (Wald and Jackson), 255.

One-band inversion (Goldschmidt and Hannah), 297.
Operators, conjugate (Miller), 359.
Operators, non-invariant (Miller), 25, 114, 275.
Oscillations, mechanical self-excited (Minorsky), 308.

Parallel flows (Lin), 316.
Penicillium notatum (Waksman and Reilly), 99.
Pericentric inversion (Giles), 1.
Permutable periodic transformations (Smith), 105.
Polyploids (Atwood), 69.
Populations, yeast (Spiegelman, Lindegren and Hedgecock), 13.
Position effect (Ephrussi and Sutton), 183; (Stern and Heidenthal), 197.
Projective collineation group (Mendelsohn), 279.
Projective space of *n* dimensions (Chern), 95.

Rabbit (Sawin, Griffith and Stuart), 217.
Radiation, ultraviolet (Luria), 392.
Rat (Castle and King), 79.
Rational diophantine analysis (Bell), 355.
Rats, waltzing in (Whittinghill), 221; (Castle), 226.
Rattus Norvegicus (Castle and King), 79.
Reciprocal chiasmata (Cooper), 50.
Red stars (Haro), 247.
REED, I. S. On the Solution of a General Transform, 169.
REILLY, H. C. See Waksman, S. A., 99.
Respiratory rates (Ball and Boell), 45.
Ring, group (Brauer), 109.
Rings, normed vector (Hille), 58.

SAWIN, P. B., GRIFFITH, M. A., AND STUART, C. A. Genetic Linkage of Blood Types in the Rabbit, 217.
Scale curves (Kasner and DeCicco), 162, 211.
SEGAL, I. E. The Span of the Translations of a Function in a Lebesgue Space, 165.
Segregation, directed (MacKnight and Cooper), 384.
Semi-groups (Hille), 58.
Serine (Tatum and Bonner), 30.
Serological methods (Sturtevant), 176.
Sets of conjugate operators (Miller), 359.
Sex chromosome (Cooper), 50.
Sex chromosomes of *Drosophila miranda* (MacKnight and Cooper), 384.
Sexual isolation in Drosophila (Dobzhansky and Mayr), 238; (Dobzhansky and Streisinger), 340; (Dobzhansky), 335.
SHAPLEY, H. Galactic and Extragalactic Studies, XVII. Revision of the

Distances of Thirty High-Latitude Globular Clusters, 61.
Shoot tips (Ball and Boell), 45.
Smith, P. A. Permutable Periodic Transformations, 105.
Sinnott, E. W., and Bloch, R. Visible Expression of Cytoplasmic Pattern in the Differentiation of Xylem Strands, 387.
Soper's epidemic curve (Wilson and Worcester), 37, 264.
Space, Lebesgue (Segal), 165.
Space, projective (Chern), 95.
Space-time (Birkhoff), 324.
Sparrow, A. H. X-Ray Sensitivity Changes in Meiotic Chromosomes and the Nucleic Acid Cycle, 147.
Specific abnormality (Dunn and Gluecksohn-Schoenheimer), 173.
Specific mutations (Sturtevant), 176.
Specificity of Guanase (Hitchings and Falco), 294.
Specificity, strain (Waksman and Reilly), 99.
Sphere bundles (Chern), 269.
Spiegelman, S., Lindegren, C. C., and Hedgecock, L. Mechanism of Enzymatic Adaptation in Genetically Controlled Yeast Populations, 13.
Spiegelman, S. See Lindegren, C. C., 346.
Spontaneous mutation (Goldschmidt), 297.
Spore discharge mechanism of common ferns (King), 155.
Stability of two-dimensional parallel flows (Lin), 316.
Stadler, L. J. The Effect of X-Rays upon Dominant Mutation in Maize, 123.
Stars, red (Haro), 247.
Steinberg, A. G. Studies on the Development of the Eye: Evidence that the Lobe^{2}, Lobe^{4}, Lobe^{5} and Eyeless^{2} Mutants of *Drosophila melanogaster* Develop in a Manner Similar to Bar, 5.
Stern, C., and Heidenthal, G. Materials for the Study of the Position Effect of Normal and Mutant Genes, 197.
Strain specificity (Waksman and Reilly), 99.
Strands, Xylem (Sinnott and Bloch), 387.
Stuart, C. A. See Sawin, P. B., 217.
Sturtevant, A. H. Can Specific Mutations Be Induced by Serological Methods?, 176.
Sumner, F. B. Vision and Guanine Production in Fishes, 285.
Susceptibles, no accession of (Wilson and Worcester), 264.
Sutton, E. See Ephrussi, B., 183.
Synapsis of sex chromosomes of *Drosophila miranda* (MacKnight and Cooper), 384.
Systems, continuous dynamic (Ingram), 370.

Tatum, E. L., and Bonner, D. Indole and Serine in the Biosynthesis and Breakdown of Tryptophane, 30.
Tatum, E. L. See Gray, C. H., 403.
Tension of an elastic cylinder (Murnaghan), 382.
Tensor, distortion (Thomas), 140.
Tetraploid hybrids (Bergner), 302.
Thomas, T. Y. Remark on a Distortion Tensor for Elastic Displacements, 140.
Thornton, H. K. See Burkholder, P. R., 250.
Thurstone, L. L. Graphical Method of Factoring the Correlation Matrix, 129.
Tissues, maturing (Ball and Boell), 45.
Transform, general (Reed), 169.
Transforms, Laplace (Chern), 95.
Translations, span of (Segal), 165.
Transformations, permutable periodic (Smith), 105.
Trifolium repens (Atwood), 69.
Trinomial congruences (Vandiver), 368.
Tropaeolum majus (Ball and Boell), 45.
Tryptophane, breakdown of (Tatum and Bonner), 30.

Uhler, H. S. First Proof that the Mersenne Number M_{157} is Composite, 314.
——. Errata, 24.
Ultraviolet radiation (Luria), 392.

Vandiver, H. S. Some Theorems in Finite Field Theory with Applications to Fermat's Last Theorem, 362.
——. On Trinomial Congruences and Fermat's Last Theorem, 368.
Varieties, higher dimensional (Chern), 95.
Virus-resistant mutants (Anderson), 396.
Vision (Sumner), 285.

Wald, G., and Jackson, B. Activity and Nutritional Deprivation, 255.
Waltzing in rats (Whittinghill), 221; (Castle), 226.
Waksman, S. A., and Reilly, H. C. Strain Specificity and Production of Antibiotic Substances, III. *Penicillium notatum*-Chrysogenum Group, 99.
Weyl, H. Comparsion of a Degenerate Form of Einstein's with Birkhoff's Theory of Gravitation, 205.

WHITTINGHILL, M. Concerning Linkage of Waltzing in Rats, 221.
WILSON, E. B., AND WORCESTER, J. A Second Approximation to Soper's Epidemic Curve, 37.
——. The Epidemic Curve with no Accession of Susceptibles, 264.
WORCESTER, J. See Wilson, E. B., 37, 264.
X-ray inhibition (Marshak and Bradley), 231.
X-rays (Bless), 118; (Stadler), 123; (Sparrow), 147; (Gray and Tatum), 403.
Xylem strands (Sinnott and Bloch), 387;
Yeast populations (Spiegelman, Lindegren and Hedgecock), 13.

I. A. R. I. 75.

IMPERIAL AGRICULTURAL RESEARCH
INSTITUTE LIBRARY
NEW DELHI.

Date of issue.	Date of issue.	Date of issue.
14-6-56		
7-4-58		
7.2.60		
18.8.1961		
18.12.1962		
24.7.63		
20.9.66		
27 DEC 1969		
28 FEB 1970		